

Research Article

In Silico Biology of H1N1: Molecular Modelling of Novel Receptors and Docking Studies of Inhibitors to Reveal New Insight in Flu Treatment

Deepak Kumar Behera,¹ Pabitra Mohan Behera,² Laxmikanta Acharya,²
Anshuman Dixit,³ and Payodhar Padhi¹

¹Hi-Tech Research & Development Centre, Konark Institute of Science and Technology, Techno Park, Jatni, Bhubaneswar, Odisha, PIN-752050, India

²Centre of Biotechnology, Siksha O Anusandhan University, Bhubaneswar, Odisha, PIN-751030, India

³Department of Translational Research & Technology Development, Institute of Life Sciences, Nalco Square, Bhubaneswar, Odisha, PIN-751023, India

Correspondence should be addressed to Payodhar Padhi, payodharpadhi@gmail.com

Received 12 March 2012; Revised 16 May 2012; Accepted 11 June 2012

Academic Editor: Nick V. Grishin

Copyright © 2012 Deepak Kumar Behera et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Influenza is an infectious disease caused by RNA viruses of the family Orthomyxoviridae. The new influenza H1N1 viral strain has emerged by the genetic combination of genes from human, pig, and bird's H1N1 virus. The influenza virus is roughly spherical and is enveloped by a lipid membrane. There are two glycoproteins in this lipid membrane; namely, hemagglutinin (HA) which helps in attachment of the viral strain on the host cell surface and neuraminidase (NA) that is responsible for initiation of viral infection. We have developed homology models of both Hemagglutinin and Neuraminidase receptors from H1N1 strains in eastern India. The docking studies of B-Sialic acid and O-Sialic acid in the optimized and energy-minimized homology models show important H-bonding interactions with ALA142, ASP230, GLN231, GLU232, and THR141. This information can be used for structure-based and pharmacophore-based new drug design. We have also calculated ADME properties (Human Oral Absorption (HOA) and % HOA) for Oseltamivir which have been subject of debate for long.

1. Introduction

The H1N1 virus particle is about 80–120 nanometers in diameter and roughly spherical [1, 2]. It is made up of a viral envelope containing two main types of glycoproteins hemagglutinin (HA) and neuraminidase (NA), wrapped around a central core [3, 4] that contains a single-stranded viral RNA. The eight single (nonpaired) RNA strands encode for eleven proteins HA, NA, NP, M1, M2, NS1, NEP, PA, PB1, PB1-F2, and PB2. HA encodes hemagglutinin while NA encodes neuraminidase. The influenza A virus can be further classified into subtypes by serological reactivity of its surface glycoprotein antigens. H1N1 is a serotype of influenza A virus that commonly causes swine flu in humans [5–8]. During infection, the influenza virus is attached to the cell receptor

(sialic acid) through HA. The HA plays an important role in the release of the viral RNA into the cell, by causing fusion of viral and cellular membranes [9]. Once the (-) strand influenza viral RNAs enter the nucleus, they serve as templates for the synthesis of mRNAs by RNA-dependent RNA polymerase [10]. The new (-) strand viral RNAs produced in the cell nucleus are exported to the cytoplasm and are joined with the viral proteins PA, PB1, PB2, and NP. M1 protein binds to the membrane on which HA, NA, and M2 have been inserted. The assembly and precise packaging of both viral RNAs and viral proteins form the new virions [11]. These are produced by budding and are attached to sialic acid receptors on the cell surface and wait for the viral neuraminidase activity that removes sialic acids from the cell surface and release them.

It is obvious that if the functionality of viral neuraminidase is rendered ineffective by any means the virions will fail to be released from infected host and further infect fresh hosts. Several neuraminidase inhibitors such as Oseltamivir and Zanamivir have been designed to work by blocking the function of the viral neuraminidase protein. Neuraminidase inhibitor treatment limits the severity and spread of viral infections. Again, if the cell receptor (sialic acid) that helps in viral attachment to the host cells is altered the virions will fail to infect. There are a number of chemically different forms of sialic acids which can be used in this regard [9, 12]. In our *in silico* study, we present molecular modelling of novel neuraminidase and hemagglutinin receptors and interaction of sialic acid types on hemagglutinin receptors to reveal novel insights for structure-based and pharmacophore-based drug design for the development of novel therapeutics.

2. Materials and Methods

2.1. Materials. The sequences for the nucleocapsid protein that is neuraminidase (NA) and hemagglutinin (HA) were taken from flu.gov database of NCBI [13]. The NCBI influenza virus sequence database contains nucleotide sequences as well as protein sequences and their encoding regions derived from the nucleotide sequences. The neuraminidase protein sequences were retrieved by putting the (keywords Type: A, Host: Human, Country/Region: India, Protein: NA, Subtype: H1 and N1, Sequence type: Protein.) We got a list of 350 protein sequence of NA of different regions of India, from which sequences of eastern India were considered for research. Similar search was performed for hemagglutinin (HA) and a list of 95 protein sequences from different regions of India of which sequence of Eastern India were considered. About 15 sequences of Neuraminidase and 95 protein sequences of hemagglutinin were downloaded in FASTA format for analysis. Prior to model development, the sequences were analyzed in BioEdit [14]. It was found that all sequences were approximately of same length. The longest sequence of NA with accession id ADD85917 was selected for 3D model development that contains 453 amino acid residues with molecular weight 49654.19 Daltons. Similarly, the longest sequence of HA with accession ID ADD85911 was selected for model development that contains 335 amino acid residues with molecular weight of 37140.01 Daltons.

2.2. Methods

2.2.1. Molecular Modelling. Molecular modelling of novel HA and NA receptor proteins was performed using modelling server (I-TASSER [15]) and standalone software (Modeller 9.9 [16–19]). For modelling in Modeller 9.9, suitable templates were searched with BLAST-P [20] against the PDB [21]. Six templates (for NA) and ten templates (for HA) were considered for modeling on the basis of query coverage and *E*-value. The alignment files of the target protein sequence with the suitable templates were generated and ten models were generated for each protein (Please see supplementary material for more information).

2.2.2. Molecular Docking. The Oseltamivir and Sialic acids (B-Sialic acid and O-Sialic acid) were obtained from PubChem [22] in SDF format and converted to PDB format for docking purpose. Schrodinger [23, 24] software was used for flexible docking. The receptors were prepared by assigning bond orders, adding hydrogens, setting proper ionization states of residues, capping the termini, and so forth. The receptors were then refined with H-bond assignment (water orientations, at neutral pH), and energy was minimized with OPLS 2005 force field. A grid for the protein was generated by using site around the centroids of selected residues. The ligands were prepared in ligprep with the following parameters force field: OPLS2005, ionization at target pH: 7.0 ± 2.0 , generate tautomers and stereo isomers (generate all combinations of specified chiralities and determine chiralities from 3D structure) with at most 32 ligands to be generated. Finally, the ligands were docked by XP (Extra Precision) method with flexibility (nitrogen inversions and ring conformations). The results were exported with XP descriptors information and adding Epik State penalties to docking score.

3. Results

3.1. The 3D Model of NA Receptor. Five 3D models were generated with different C-scores featuring the accuracy of prediction from I-TASSER. The best model was chosen on the basis of highest C-score. One model of NA was selected from those generated in Modeller 9.9 on the basis of different scores (molpdf: 15581.55371, DOPE score: -44868.22266 and GA341 score: 1.00000).

3.2. The 3D Model of HA Receptor. Best model was selected from among those generated by I-TASSER with C-score = 1.47, estimated accuracy (TM-score) = 0.92 ± 0.06 . Similarly, a model of HA was selected from amongst generated by Modeller 9.9 based on various scores (molpdf: 15373.00293, DOPE score: -33933.94141 and GA341 score: 1.00000).

3.3. Evaluation of the Models. PROCHECK checks the stereochemical quality of a protein structure, producing a number of PostScript plots analyzing its overall geometry. All the models were evaluated in Procheck_NT standalone version. It produced 10 postscript files featuring Ramachandran plots, Chi1-Chi2 plots, main chain parameters, side chain parameters, residue properties, main chain bond lengths, main chain bond angles, RMS distances from planarity, and distorted geometry. The selected models were found to be satisfactory for the calculated stereochemical parameters.

3.4. Flexible Docking with Schrodinger. Docking results of Oseltamivir, beta-sialic acid, and ortho-sialic acid were imported in XP visualiser and corresponding docking poses were viewed in workspace (Figures 3 and 4). The overall dock scores (GScore) for the three ligands were -5.18 , -9.74 , and -8.65 , respectively (Table 1). The GScore is the total glide score based on ChemScore used for ranking ligand poses found in docking. The H-bond interaction among the ligand

TABLE 1: Docking of Oseltamivir, Beta-Sialic acid, and O-Sialic acid from XP method.

Sl. no.	Ligands	GScore	Lipophilic	H-bond	Electrostatic	Low MW
1	Oseltamivir	-5.18	-1.10	-1.50	-2.67	-0.46
2	B-Sialic acid	-9.74	-0.92	-6.77	-2.00	-0.47
3	O-Sialic acid	-8.65	-0.93	-5.73	-2.00	-0.47

Lipophilic: ChemScore lipophilic pair term and fraction of a protein-ligand vdW energy.

Hbond: ChemScore H-bond pair term.

Electrostatic: electrostatic rewards.

Low MW: Rewards for ligand with low molecular weight.

TABLE 2: Model evaluation statistics of NA and HA by PROCHECK_NT standalone version.

Protein	Programs	[A, B, L]	[a, b, l, p]	[~a, ~b, ~l, ~p]	Disallowed region
NA	Modeller 9.9	354 (91.20%)	30 (07.70%)	3 (0.80%)	1 (0.30%)
	I-TASSER	324 (83.50%)	57 (14.70%)	4 (1.00%)	3 (0.80%)
HA	Modeller 9.9	264 (90.10%)	27 (9.20%)	2 (0.70%)	0 (0.00%)
	I-TASSER	228 (77.80%)	60 (20.50%)	3 (1.00%)	2 (0.70%)

and the receptor residues was analyzed. Detailed information about the docking (reward and penalty) was reported from XP visualiser.

4. Discussions

The structures of NA receptor modelled by Modeller 9.9 and I-TASSER reported the amino acids percentage in the favourable region as 91.20% and 83.50%, respectively, with amino acids percentage in the disallowed region as 0.30% and 0.80%, respectively, similarly the structures of HA receptor modelled by Modeller 9.9 and I-TASSER reported the amino acids percentage in the favorable region as 90.10% and 77.80%, respectively, with amino acids percentage in the disallowed region as 0.00% and 0.70%, respectively (Table 2).

All these models were evaluated by PROCHECK_NT [25] standalone version which indicates that the model build by Modeller 9.9 shows the highest percentage of amino acids in the allowed region. Further, these structures were validated with PROSA 2003 standalone version which produced Z-Score in the range -5 to -7 . Hence, these models were considered as best models and are represented in Figures 1 and 2. The best model of NA has 3 α -helices, 25 β -sheets, and 15 loops while the best model of HA has with 5 α -helices, 24 β -sheets, and 11 loops.

4.1. Docking of Oseltamivir. Oseltamivir ($C_{16}H_{28}N_2O_4$) with four hydrogen-bond acceptors and two hydrogen-bond donors showed 4 hydrogen-bonds on the NA receptor which was quite interesting. This docking model can be used for generation of active site-based pharmacophore for virtual screening of chemical databases.

4.2. Docking of B-Sialic Acid and O-Sialic Acid. Beta-sialic acid ($C_{11}H_{19}NO_9$) with nine hydrogen-bond acceptors and seven hydrogen-bond donors showed 8 hydrogen-bonds on the HA receptor while O-Sialic Acid ($C_{11}H_{19}NO_9$) with nine

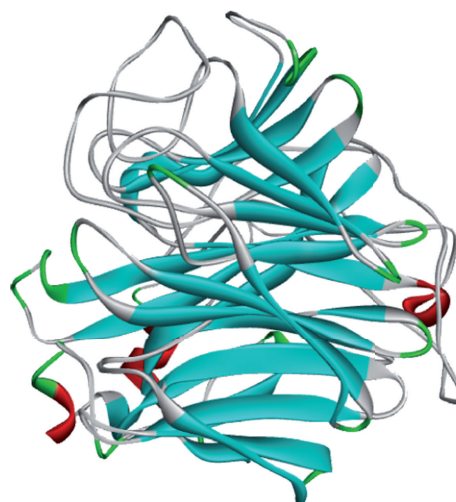


FIGURE 1: 3D representation of NA best model.

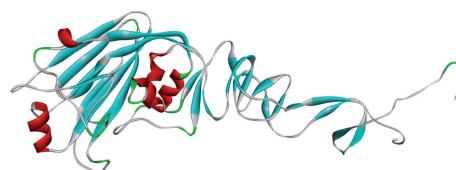


FIGURE 2: 3D representation of HA best model.

hydrogen-bond acceptors and seven hydrogen-bond donors showed 8 hydrogen-bonds on the HA receptor.

The docking of Oseltamivir on Neuraminidase shows H-bond interaction with residues ARG109, ARG143, and GLU219. The docking of beta-sialic acid on Hemagglutinin shows Hydrogen-bond interaction with residues ALA142, ASP230, GLN231, GLU232, and THR141. Again the docking of O-sialic acid on Hemagglutinin shows hydrogen-bond interaction with residues ALA142, ASP230, GLN231, GLU232, and THR141. This residue information (Table 3)

TABLE 3: Hydrogen-bond interaction parameters for each compound with NA and HA receptors.

Sl. no.	Residue involved	Ligand atom involved	H-bond distance (Å)
Oseltamivir on NA receptor			
1	ARG109	O	1.832
2	ARG143	O	2.280
3	ARG143	O	2.407
4	GLU219	N	1.760
B-Sialic on HA receptor			
1	ALA142	O	2.137
2	ASP230	O	2.047
3	GLN231	O	2.032
4	GLN231	O	1.965
5	GLN231	O	1.910
6	GLU232	O	2.182
7	THR141	O	2.224
8	THR141	O	1.796
O-Sialic on HA receptor			
1	ALA142	O	2.136
2	ASP230	O	1.966
3	ASP230	O	1.928
4	GLN231	O	2.073
5	GLN231	O	1.995
6	GLU232	O	2.217
7	THR141	O	2.215
8	THR141	O	1.830

TABLE 4: Important ADME properties of Oseltamivir conformers as reported from Schrodinger.

Sl. no.	Molecule	QPlogS	QPlogBB	QPlogKp	HOA	Percent HOA
1	Oseltamivir	-1.927	-0.921	-5.303	3	66.024
2	Oseltamivir	-2.389	-0.809	-5.120	3	70.229
3	Oseltamivir	-2.199	-1.021	-5.420	3	65.346
4	Oseltamivir	-2.017	-0.780	-5.121	3	69.130
5	Oseltamivir	-1.763	-0.782	-5.110	3	68.065
6	Oseltamivir	-2.231	-0.769	-5.086	3	70.294
7	Oseltamivir	-1.981	-0.867	-5.222	3	67.360
8	Oseltamivir	-1.993	-0.907	-5.304	3	66.440
9	Oseltamivir	-1.984	-0.929	-5.340	3	65.989
10	Oseltamivir	-2.438	-0.788	-5.078	3	70.885
11	Oseltamivir	-2.035	-0.897	-5.285	3	66.988
12	Oseltamivir	-2.035	-0.764	-5.102	3	69.586
13	Oseltamivir	-1.870	-0.948	-5.375	3	65.114
14	Oseltamivir	-2.170	-0.781	-5.107	3	69.772
15	Oseltamivir	-1.928	-0.927	-5.312	3	65.961
16	Oseltamivir	-2.010	-0.902	-5.293	3	66.648

QPlogS: predicted aqueous solubility.

QPlogBB: predicted brain/blood partition coefficient.

QPlogKp: predicted skin permeability.

HOA: predicted qualitative human oral absorption.

PercentHOA: predicted human oral absorption on 0 to 100% scale.

can be used for further high-throughput screening of novel bioactive molecules.

4.3. QikProp Functional Assessment. QikProp is a quick, accurate, easy-to-use absorption, distribution, metabolism, and excretion (ADME) prediction program. QikProp

predicts physically significant descriptors and pharmaceutically relevant properties of organic molecules, either individually or in batches. In addition to predicting molecular properties, QikProp provides ranges for comparing a particular molecule's properties with those of 95% of known drugs. QikProp also flags 30 types of reactive functional groups

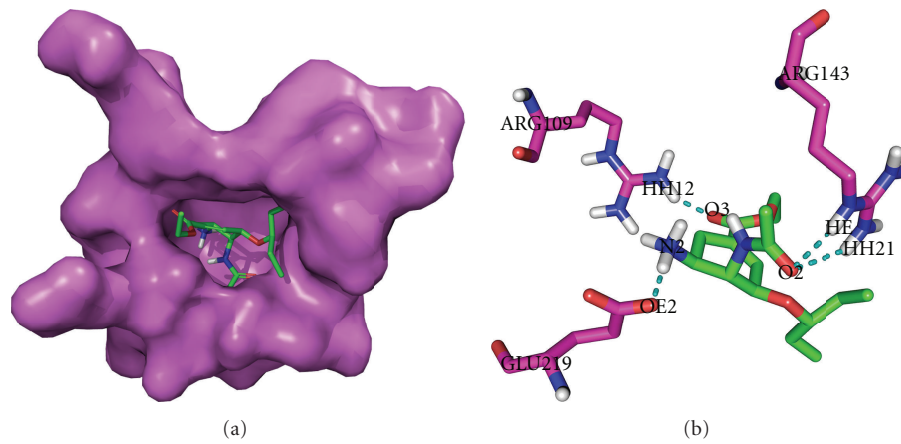


FIGURE 3: (a) Docking of Oseltamivir on the NA surface reported from Schrodinger, (b) interaction of Oseltamivir on the NA by hydrogen-bond.

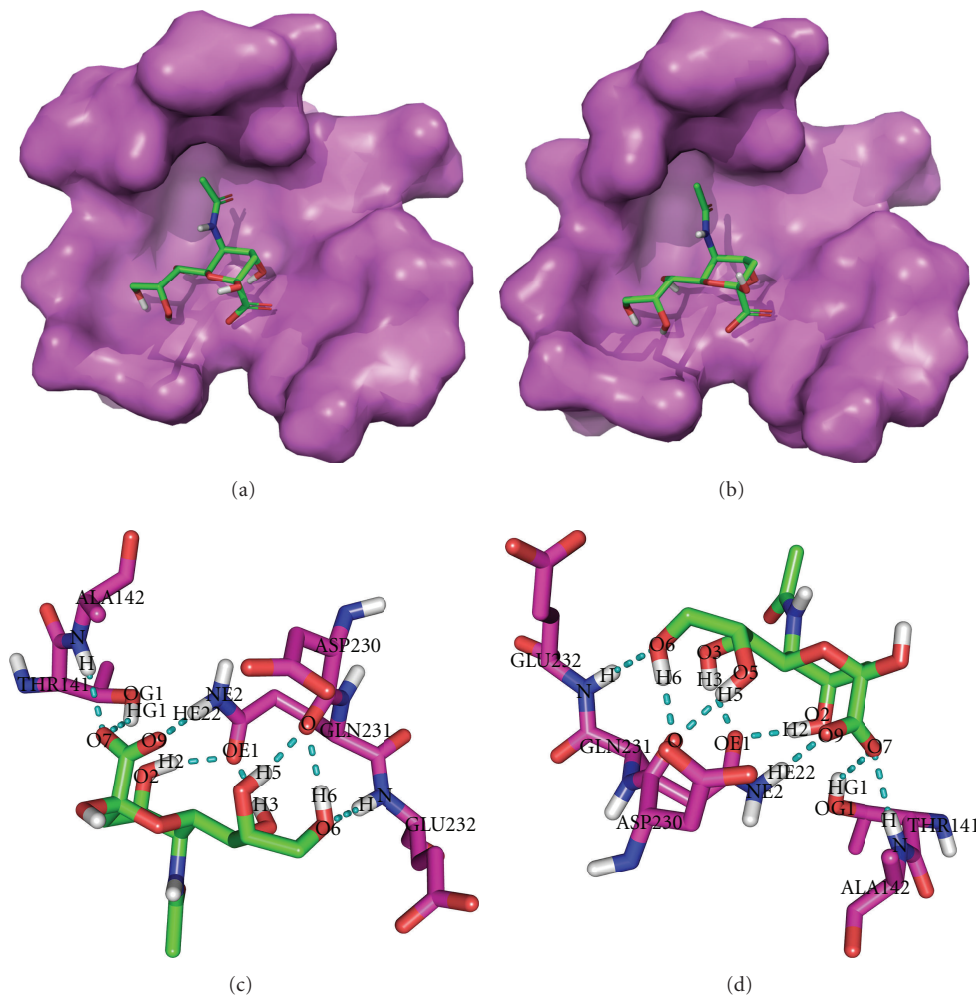


FIGURE 4: (a) Docking of Beta-Sialic acid on molecular surface of HA protein by Schrodinger, (b) docking of O-Sialic acid on molecular surface of HA protein by Schrodinger; (c) interaction of B-Sialic acid with HA protein by hydrogen-bond, and (d) interaction of O-Sialic acid with HA protein by hydrogen-bond.

that may cause false positives in high-throughput screening (HTS) assays. The range of values that cause a molecule to be flagged as dissimilar to other known drugs can be modified in the QPLimits. The 16 conformers of Oseltamivir obtained from ligprep were imported to QikProp and their ADME properties were generated using fast mode. About 51 descriptors and properties were reported of which few important are represented in Table 4. According to the Hazardous Substances Data Bank [26] the percent Human Oral Absorption is 75%. Earlier it was debated that Oseltamivir (Tamiflu) has a low rate of absorption but our predicted qualitative human oral absorption for all 16 conformers was 3 (the highest) and the percent of human oral absorption was >67.739% (average). Again, the predicted values of aqueous solubility, brain/blood partition coefficient, and skin permeability for all 16 conformers were good.

5. Conclusion

The docking studies of Oseltamivir (Tamiflu) on the optimized and energy-minimized model of NA showed some important H-bond interactions with functionally important residues. Similarly, the docking studies of B-Sialic acid and O-Sialic acid on the optimized and energy-minimized model of HA show some important H-bond interactions. This information can be used for structure-based and pharmacophore-based new drug designing for development of novel therapeutic agents for the prevention and treatment of influenza.

Acknowledgments

The authors are thankful to the Vice Chancellor of Siksha O Anusandhan University and the Chairman of Hi-Tech Group of Institutions for providing essential facilities in successful research and bringing the paper to the level of publication. The award of Rajiv Gandhi National Fellowship to P. M. Behera by University Grants Commission, India is deeply acknowledged.

References

- [1] C. Büchen-Osmond, Ed., "ICTVdB Virus Description—00.046.0.04. Influenzavirus B". *ICTVdB—The Universal Virus Database, Version 4*, vol. 142, Columbia University, New York, NY, USA, 1997.
- [2] R. A. Lamb and P. W. Chopin, "The gene structure and replication of influenza virus," *Annual Review of Biochemistry*, vol. 52, pp. 467–506, 1983.
- [3] Y. Kawaoka, Ed., *Influenza Virology: Current Topics*, Caister Academic Press, 2006.
- [4] "International Committee on Taxonomy of Viruses Index of Viruses—Orthomyxoviridae," in *ICTVdB—The Universal Virus Database, Version 4*, C. Büchen-Osmond, Ed., Columbia University, New York, NY, USA.
- [5] J. S. Malik Peiris, L. L. M. Poon, and Y. Guan, "Emergence of a novel swine-origin influenza A virus (S-OIV) H1N1 virus in humans," *Journal of Clinical Virology*, vol. 45, no. 3, pp. 169–173, 2009.
- [6] D. Gatherer, "The 2009 H1N1 influenza outbreak in its historical context," *Journal of Clinical Virology*, vol. 45, no. 3, pp. 174–178, 2009.
- [7] G. Neumann, T. Noda, and Y. Kawaoka, "Emergence and pandemic potential of swine-origin H1N1 influenza virus," *Nature*, vol. 459, no. 7249, pp. 931–939, 2009.
- [8] M. Michaelis, H. W. Doerr, and J. Cinatl, "Novel swine-origin influenza A virus in humans: another pandemic knocking at the door," *Medical Microbiology and Immunology*, vol. 198, no. 3, pp. 175–183, 2009.
- [9] C. Liu, M. C. Eichelberger, R. W. Compans, and G. M. Air, "Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding," *Journal of Virology*, vol. 69, no. 2, pp. 1099–1106, 1995.
- [10] E. Ghedin, N. A. Sengamalay, M. Shumway et al., "Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution," *Nature*, vol. 437, no. 7062, pp. 1162–1166, 2005.
- [11] Y. Muramoto, A. Takada, K. Fujii et al., "Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions," *Journal of Virology*, vol. 80, no. 5, pp. 2318–2325, 2006.
- [12] M. von Itzstein, "The war against influenza: discovery and development of sialidase inhibitors," *Nature Reviews Drug Discovery*, vol. 6, no. 12, pp. 967–974, 2007.
- [13] Y. Bao, P. Bolotov, D. Dernovoy et al., "The influenza virus resource at the National Center for Biotechnology Information," *Journal of Virology*, vol. 82, no. 2, pp. 596–601, 2008.
- [14] T. A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symposium Series*, vol. 41, pp. 95–98, 1999.
- [15] A. Roy, A. Kucukural, and Y. Zhang, "I-TASSER: a unified platform for automated protein structure and function prediction," *Nature Protocols*, vol. 5, no. 4, pp. 725–738, 2010.
- [16] N. Eswar, M. A. Marti-Renom, B. Webb et al., *Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics*, John Wiley & Sons, 2006.
- [17] M. A. Marti-Renom, A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, and A. Šali, "Comparative protein structure modeling of genes and genomes," *Annual Review of Biophysics and Biomolecular Structure*, vol. 29, pp. 291–325, 2000.
- [18] A. Sali and T. L. Blundell, "Comparative protein modelling by satisfaction of spatial restraints," *Journal of Molecular Biology*, vol. 234, no. 3, pp. 779–815, 1993.
- [19] A. Fiser, R. Kinh Gian Do, and A. Sali, "Modeling of loops in protein structures," *Protein Science*, vol. 9, no. 9, pp. 1753–1773, 2000.
- [20] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [21] The Research collaborator for Structural Bioinformatics, Protein Data Bank, <http://www.rcsb.org/pdb/home/>.
- [22] E. E. Bolton, Y. Wang, P. A. Thiessen, and S. H. Bryant, *PubChem: Integrated Platform of Small Molecules and Biological Activities. Chapter 12 in Annual Reports in Computational Chemistry*, vol. 4, American Chemical Society, Washington, DC, USA, 2008.
- [23] B. Dineshkumar, P. Vignesh Kumar, S. P. Bhuvaneshwaran, and A. Mitra, "Advanced drug designing softwares and their applications in medical research," *International Journal of Pharmacy and Pharmaceutical Sciences*, vol. 2, no. 3, pp. 16–18, 2010.
- [24] R. A. Friesner, J. L. Banks, R. B. Murphy et al., "Glide: a new approach for rapid, accurate docking and scoring 1. Method

and assessment of docking accuracy,” *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1739–1749, 2004.

- [25] R. A. Laskowski, M. W. MacArthur, and J. M. Thornton, “PROCHECK: validation of protein structure coordinates,” in *International Tables of Crystallography*, M. G. Rossmann and E. Arnold, Eds., Crystallography of Biological Macromolecules, pp. 722–725, Kluwer Academic, Dordrecht, The Netherlands, 2001.
- [26] G. C. Fonger, “Hazardous substances data bank (HSDB) as a source of environmental fate information on chemicals,” *Toxicology*, vol. 103, no. 2, pp. 137–145, 1995.