Drug and Anti-Viral Peptide Design to Inhibit the Monkeypox Virus by Restricting A36R Protein

Mohammad Mohasin Miah1, Nuzhat Tabassum1, Maliha Afroj Zinnia1 and Abul Bashar Mir Md. Khademul Islam2

1Department of Pharmacy, East West University, Dhaka, Bangladesh. 2Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka, Bangladesh.

ABSTRACT: Most recently, monkeypox virus (MPXV) has emanated as a global public health threat. Unavailability of effective medicament against MPX escalates demand for new therapeutic agent. In this study, in silico strategies were conducted to identify novel drug against the A36R protein of MPXV. The A36R protein of MPXV is responsible for the viral migration, adhesion, and vesicle trafficking to the host cell. To block the A36R protein, 4893 potential antiviral peptides (AVPs) were retrieved from DRAMP and SATPdb databases. Finally, 57 sequences were screened based on peptide filtering criteria, which were then modeled. Likewise, 31 monkeypox virus A36R protein sequences were collected from NCBI protein database to find consensus sequence and to predict 3D protein model. The refined and validated models of the A36R protein and AVP peptides were used to predict receptor-ligand interactions using DINC 2 server. Three peptides that showed best interactions were SATPdb10193, SATPdb21850, and SATPdb26811 with binding energies −8.8 to 9.8 kcal/mol. Interaction of target protein A36R with the top 3 peptides and the most probable drug (Z55287118) examined by molecular dynamic (MD) simulation. Trajectory analyses (RMSD, RMSF, SASA, and Rg) confirmed the stable nature of protein-ligand and protein-peptide complexes. This work suggests that identified top AVPs and small molecules might interfere with the function of the A36R protein of MPXV.

KEYWORDS: MPXV, A36R protein, AVPs, molecular docking, protein-peptide docking, molecular dynamics simulation, ADMET analysis

Introduction

The eradication of smallpox from human populations was one of the 20th century’s most remarkable victories in medicine. One of the grizzly incompatibilities of this achievement is the emergence of Monkeypox (MPX) with significant mortality and morbidity.1 The recent surge in human monkeypox cases outside of their normal ecological range, the possibilities for advance spread through the transmission to naive populations as half of the world’s population has no immunity against orthopoxviruses,1 and the lack of established surveillance have intensified the level of concern for this emerging zoonosis.4

The zoonosis human monkeypox (MPX) is caused by monkeypox virus (MPXV) which belongs to the genus Orthopoxvirus, family Poxviridae, and sub-family Chordopoxvirinae.

MPXV was first unearthed during an outbreak in captive monkeys at a Danish laboratory in 1958.5,6 Congo Basin (CB) and West African (WA) are the two distinct clades of MPXV.7 The CB clade primarily infects people residing in the Congo Basin.8,9 Between 1996 and 1997, more than 400 incidences were reported in Democratic Republic of the Congo (DRC) by CB clade followed by other outbreaks with hundreds of cases in the 2000s.9,11 Since 2016, DRC continues to report the most cases each year mainly in children aged below 10 years.4,5,12,13

and clustering of cases was reported in some particular areas within countries and within families.14

In West Africa, only sporadic cases were revealed between 1970 and 1981. The largest known outbreak of monkeypox virus to date by WA clade with 122 confirmed or probable cases were recorded in 17 states of Nigeria in 2017-2018.15 The characteristics of the 2017-2018 outbreak by WA clade suggest that the strains of MPXV in West Africa could prolong epidemic events, although West Africa was previously thought to be at low risk of a human monkeypox endemic.8

Human monkeypox prevalence apart from Africa was first reported in the United States in 2003.2,9,14 The recent detection of MPX cases in each parts of the earth is an unprecedented and a big concern for global health security.

The majority of the clinical manifestations of human monkeypox infection mirrors those of smallpox but differs from it epidemiologically.14,17-19 A distinctive 2-day prodrome, apparent by fever, headache, muscle aches, backache, and malaise, reveals in most patients before the development of the rash.9 Lymphadenopathy, a key distinguishable feature of human monkeypox which is not characteristic of smallpox,14,20 occurs in many victims 1 to 2 days before the inception of the rash.

MPXV is an enveloped double-stranded DNA virus with an approximate genome size of 196 858 base pairs.21 Unlike many other DNA viruses, the replication cycle of poxviruses takes

*Contributed equally.
place in the cytoplasm of infected cells rather than in the nucleus.22 Poxviruses enter cells by different mechanisms depending on whether they are intracellular mature virus (IMV) or extracellular enveloped virus (EEV). The IMV is encircled by a single membrane, which enters the cell by either direct fusion with the plasma membrane or endocytosis. By contrast, the EEV is composed of taking IMV enclosed within a second lipid membrane and forms shedding with the external membrane outside of the cell and then fusion of the IMV membrane with the host cell membrane.23 Upon fusion with the host cell membrane, virus loses membrane to release its core into the cytoplasm that is transported on microtubules for moving deeper into the cell, where transcription and translation are occurred to form immature virions (IV), which are later processed to compose IMV.23,24 IMV particles are encased by a double membrane to constitute intracellular enveloped viruses (IEVs) that are further moved to the cell periphery on microtubules. The external IEV membrane attaches with the plasma membrane to expose a cell-associated extracellular virus (CEV) at the cell surface. Polymerization of an actin tail underneath the CEV drives the virus into the surrounding cell, or the virus is released as an EEV.23,25,26 According to Münter et al,27 IEV transportation to the cell periphery necessitates the microtubule motor kinesin-1, which binds the integral viral membrane protein A36R. Surprisingly, the same protein gets localized beneath the cell-associated extracellular virus (CEV), detaches from kinesin-1, and shifts to signaling actin polymerization following tyrosine phosphorylation by Src and Abl family kinases.27 Phosphorylation on Y112 and Y132 residues in A36R generates binding sites for the Src homology 2 (SH2)—domains of Nck and Grb2 adaptor proteins, respectively, that functionally attach to the actin binding proteins WIP and N-WASP.24 N-WASP stimulates the nucleation of actin filaments making connections with actin and the Arp2/3 complex, a seven-subunit complex that drives dynamic actin assembly during migration, adhesion, or vesicle trafficking.28,30

These studies make it prominent that blocking the A36R protein of poxviruses prevents the actin polymerization thus warding off the viral migration, adhesion, and vesicle trafficking to the host cell. All the viruses of the genus Orthopoxvirus encompass A36R protein functioning in the same manner.24 Therefore, to prevent the actin polymerization of the host cell the A36R protein of MPXV is thought to be a more promising drug target.

Peptides have evolved as extremely selective and potent signaling molecules that bind to specified cell surface receptors, or ion channels to exert powerful physiological effects. Their potential pharmacological profile and inherent features represent an exceptional starting point for the development of advanced therapeutics.31 As a drug, peptides are prominent for their exorbitant affinity, specificity to a respective target and efficacy, at the same time, safety and tolerability in human that estimate assorted supremacy over small molecules.32-35 In 2011, US Food and Drug Administration (FDA) corroborated peptide-based drug against the Hepatitis C virus and multiple peptides are experiencing phase trials targeting Influenza virus, Hepatitis B and D virus. Peptide therapeutics against Chikungunya virus, Zika virus, and Dengue virus are exhibiting antiviral activity.36

There was no authorized treatment against human MPXV before 2019. Dryvax, a smallpox vaccine, was practiced for both smallpox and monkeypox treatment,37 developed various influential adverse effects in vaccinated people and persons in contact with the vaccine.38,39 Later information revealed that the prime mode of protection against MPX managed by the existing non-attenuated smallpox immunization is interfered by Abs.40 Considering drawbacks of the current treatment, efforts were made to identify plausible small molecule as well peptide-based therapeutic agents against the A36R protein of MPXV conducting virtual screening strategies combined with molecular docking.

Materials and Methods
Methodological overview is demonstrated in Figure 1.

Peptide Preparation
Peptide retrieving
Potential 3459 antiviral peptides (AVPs) were retrieved from the SATPdb v1.0.41 Among the 3459 AVPs, 331 sequences were excluded as those were in modified forms. Similarly, 2042 AVPs were extracted from the data repository of antimicrobial peptides (DRAMP v3.0).42 From 3128 and 2042 AVP sequences, unique AVP set was created using Venn diagram43,44 for further analysis.

Peptide screening based on physicochemical properties
Physical and chemical parameters of the curated AVP sequences were determined using Expasy's ProtParam45 in Linux operating system. Peptide filtering criteria set in accordance with physicochemical frameworks are shown in Table 1. Accordingly, the AVPs in the rage of 5 to 22 amino acids in length were selected. The peptides having instability index value over 40, and in vivo or in vitro half-lives less than 16h, were removed. In addition, peptides having aliphatic index value ranging from 71.13 to 143.54 were kept in record. Also, positive grand average of hydropathicity (GRAVY) value containing AVPs were eliminated.

In silico structural modeling and validation
From amino acid sequences, 3 dimensional (3D) structures were modeled using the bio tools PEP-FOLD 3,53 and predicted structures were validated using structural analysis and verification server (SAVES v6.0) and protein structure analysis (ProSA-web) program.54
Energy minimization of the modeled structure

From the 3D structures, water molecules were removed and hydrogen atoms were added using AutoDock tools (Version 1.5.6). The energy of the final cleaned structures was minimized by implementing Swiss-PdbViewer v4.1.0.

Antigenicity and allergenicity prediction

To determine allergenicity like antigenicity and immunogenicity of AVPs, AllergenFP v1.0 and AllerTOP v2.0 tools were used. Toxicity of the peptides were determined using ToxinPred bio tools.

Protein Preparation

Retrieval of A36R protein sequences and multiple sequence alignment

The monkeypox A36R protein sequences were retrieved from the NCBI protein database. Molecular Evolutionary Genetics Analysis (MEGA v11) software was applied to perform multiple sequence alignment (MSA) of the protein sequences. MUSCLE algorithm with UPGMA cluster method was used for MSA.

3D structure prediction of the target receptor

The 3D structure of the monkeypox virus A36R protein (NCBI GenBank ID: QJJQ40287.1) was developed using two different bioinformatics tools, namely, Contact-guided Iterative Threading ASSEmbly Refinement (C-1-TASSER) and RaptorX as no crystal structure of the A36R protein was found in the protein data bank (PDB) and UniProt. Homology modeling of A36R protein was also conducted and the sequence data of A36R protein was compared with another protein having both a known sequence and a known structure (PDB hit: 1sl6A). The highest similarity identity to a known protein template is 16.67%, a value less than 40%, so A36R protein structure could not be built using homology modeling. Therefore, ab initio 3D structure was modeled using C-1-TASSER and RaptorX servers.

Refinement and validation of the modeled 3D structure of the target receptor

The predicted PDB format of the monkeypox virus A36R protein was refined exploiting 3Drefine bio tools, and the refined structures were further rectified taking advantage of GalaxyWEB server.

Validation of the modeled 3D structure of the target receptor

The refined PDB model of the A36R protein was validated by structural analysis and verification server (SAVES v6.0) and MolProbity, ERRAT, Ramachandran Plot, and overall G-factors from SAVES v6.0 server were used to measure the maximum quality of the predicted structures.

Figure 1. Methodological overview.
Energy minimization of the target receptor

Water molecules of the refined A36R protein were deleted to avoid untoward molecular interaction, and hydrogen atoms were attached using MGL Tools (Version 1.5.6). The energy minimization of A36R protein was performed using Avogadro software with the support of MMFF94 force field, setting up 500 steps, steepest descent algorithm and 10e-7 convergence, respectively. The minimized protein 3D structure was used for molecular docking and dynamics simulation.

Prediction of active sites

The ligand-binding pocket, ie, inhibitory site of the A36R protein, was determined applying the Computed Atlas for Surface Topography of Proteins (CASTp 3.0) (Figure 2A), DoGSiteScorer (Figure 2B), and DeepSite (Figure 2C), respectively. AutoDock Vina was also used to generate the grid box size 30, 30, 30 for x, y, z axes, grid center 38.4 × 54.2 × 0.2, respectively, and 0.375 Å spacing.

Molecular docking

The PDB structure of the A36R protein as the target receptor and the peptides as the ligands were provided as input in the DINC 2.0 server, which is a parallelized meta-docking method for the incremental docking of large ligands (using AutoDock Vina).

Virtual screening of A36R protein

TACC, a virtual drug discovery portal, was used to conduct virtual screening. The AutoDock Vina application recommended that our target structure A36R be imported in PDB format and subsequently converted to PDBQT format. Aside from the target structure, the center and size of the search space, as well as the parameters of the designated virtual libraries for screening from the set of accessible compound libraries, were also provided for the initiation of the virtual screening experiment. The grid box’s dimensions and coordinates were center x = -38.4, center y = 54.2, center z = 0.2, size x = 30, size y = 30, and size z = 30. The exhaustiveness was adjusted to 10 for molecular docking simulations. In high-throughput virtual screening of the A36R, the ZINC-in-trials (9270 molecules), ZINC-fragments (546,003 fragments), and Enamine-HTSC (3,467,770 molecules) libraries were employed. The ZINC library in-tails, on the other hand, contains medications that are now being tested in clinical trials.

Pharmacokinetic profile (ADMET) evaluation of the hit compounds from virtual screening

A preliminary in silico pharmacokinetic study was conducted with ADMETlab 2.0, which includes absorption (Papp Caco-2 permeability, human intestinal absorption), distribution (plasma protein binding [PPB], blood-brain barrier [BBB]), metabolism (P450 CYP3A4 inhibitor, P450 CYP2C9 inhibitor), elimination (half-life time [T1/2], clearance rate [CL]), and toxicity (hERG toxicity, AMES toxicity, hepatotoxicity, respiratory toxicity).

Molecular dynamics (MD) simulation

To appraise the ligand’s effectiveness in suppressing A36R of the monkeypox virus, MD simulations of the protein–ligand (drug molecule) complex were undertaken. We conducted the MD simulation as described in our earlier work and few other studies. GROMACS (v2021) with the company of the CHARMM General Force Field (CGenFF) following energy minimization was used to accomplish the MD simulation. To develop the system, an in-house python script was used which used Chimera dock prep functionality to cleanse the protein

### Table 1. Peptide filtering criteria based on length, instability index, aliphatic index, half-life, and grand average of hydropathicity (GRAVY).

| CRITERIA       | CUTOFF | REMARKS                                                                 | REFERENCES |
|---------------|--------|--------------------------------------------------------------------------|------------|
| Length        | 5 to 22| Very short peptides are less likely to have a unique sequence, and they are more susceptible to interferences. On the other hand, lengthy peptides are not desirable due to their hydrophobicity and to issues regarding their synthesis and purification. Typically, peptides ranging from 5 to around 22 amino acids are best suited for quantification purpose and interferences. |
| Instability Index (II) | <40    | A protein whose instability index is smaller than 40 is predicted as stable. |
| Aliphatic Index (AI) | 71.13 to 143.54 | A high aliphatic index indicates that a protein is thermo-stable over a wide temperature range. |
| Half-life     | >16 h  | Rogers et al classified proteins into sets (that were used by Guruprasad et al 1990, and others) as stable if the in vivo half-lives were greater than 16 h or unstable if the in vivo half-lives were less than 5 h. |
| GRAVY         | Negative| Positive GRAVY values indicate hydrophobic; negative values mean hydrophilic. Low GRAVY range indicates the possibility of being a globular (hydrophilic) protein rather than membranous (hydrophobic). |

46,47 48 49 50 51,52
and to attach missing residues exploiting the Dunbrack rotamer library. Hydrogens were physically put into the protein employing the GROMACS pdb2gmx function, whereas hydrogens were manually placed into the ligand using custom force field settings. The GROMACS tool was imposed to create the protein topology for A36R, and parameterization of the ligand topology was achieved by applying the LigParGen program. The LigParGen bio tool accomplished optimization using the CGenFF, and the charge model (CM1A) parameters were also developed. The system was put in the center of a cubic box of a TIP3P water model with a minimum distance of 1.0 nm between the wall and any component of the protein set up before the start of the experiment. Normalization of the solvated system was done by adding Na\(^+\) (sodium) and Cl\(^-\) (chloride) to a 0.15 M ionic strength aqueous solution. Canonical (NVT) and constant-pressure, constant-temperature (NPT) ensemble were used to perform system equilibration for 100 picoseconds (ps). The trajectories were also constructed in every 2 femtoseconds (fs) and recorded every 1 ps. The minimized system was progressively heated to appropriate temperatures for 100 ps within an isothermal ensemble by soft coupling with the Berendsen thermostat (NVT). To avoid drift away from the protein during equilibration, ligand location limitations were applied before running NVT simulations. All bonds were restricted using the LINCS algorithm. Periodic boundary conditions (PBC) were applied with a constant number of particles in the system, constant pressure, and constant temperature simulation requirements (NPT). The NPT ensemble was executed at a temperature of 300 K. The system was combined with a Parrinello–Rahman barostat for equilibrating at 1 bar pressure for 100 ps. The Particle Mesh Ewald (PME) method was used to process the electrostatic interactions. With a cutoff of 1.0 nm, the short-range van der Waals cutoff (rvdW) interactions were calculated.

Production simulations with limitations were run for 10 ns. The LINCS algorithm was taken to limit all bond lengths. DELL workstation with Intel Xenon processor (10C, 20T), 256 GB RAM; running Fedora (v31) Linux operating system was employed to run MD. The root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), hydrogen bond (H-bond), solvent accessible surface area (SASA) (using van der Waals Volumes and Radii), and distance analysis were computed using Gromacs. Trajectory plots and figures were developed to exploit an in-house python script with Matplotlib and NumPy library as well as implementing R version 3.6.3 and Peptides library.

Result

Peptide selection

We obtained 4893 AVP sequences for scrutiny by Expasy’s ProtParam (Proteomics Protoc. Handb., 2005), after excluding 277 common AVP set from either database and 331 revised versions from SATPdb. Among 4893 AVPs, 57 sequences exceeded the selection benchmarks of length, instability index, aliphatic index, half-life, and grand average of hydrophobicity, respectively (Supplementary Table A.1). This selection procedure had a high resemblance with the peptide filtering technology reported by Chang and Yang, 2013.

Validation analysis of the predicted 3D structures

Five crystal structures were received against each insertion of sequences and overall 285 3D models were attained from PEPFOLD 3 bio tools. After implementing 5 models individually to SAVES v6.0 server and ProSA-web program, 1 model was validated because of showing maximum percentage of residues in the most favored region of Ramachandran Plot and

![Image of 3D structure of A36R and its predicted active site: (A) Showing the inhibitory site by using CASTp, (B) DeepSite, and (C) DoGSiteScorer. CASTp indicates Computed Atlas for Surface Topography of Proteins.](image-url)
exhibited significant ERRAT score, weighty G-factors and notable ProSA-web Z score. Thus, 57 models were finalized those revealed the highest ERRAT score, greater core region of Ramachandran plot, and appropriate G-factors and Z-score, respectively (Table 2).

**Multiple sequence alignment analysis of central and western strain of monkeypox virus (MPV) A36R protein**

Total 31 monkeypox virus A36R protein sequences of both West African and Central African strain were retrieved from NCBI protein database. The sequences comprised 168 amino acids in length. MSA (Supplementary Fig. A.1) execution of 31 MPXV A36R protein sequences exhibited mutation in the 3 sequences at 1 point (glycine instead of arginine) compared with others and these mutated sequences carrying 11 patients were isolated in Central Africa and named Zaire 96-I-96 strain or Monkeypox virus Z1 strain. Therefore, we excluded mutation containing 3 sequences (Accession: AAL40604.1; NP_536573.1 and CAC69954.1) and selected West African strain as our final A36R protein sequence. Finally, NCBI GenBank ID holding QJQ40287.1 sequence from West African strain was chosen to predict 3D structure.

**Structural modeling and validation analysis of monkeypox virus (MPV) A36R protein**

Predicted 5 models of C-I-TASSER generated 25 refined 3D structures after employing 3Drefine bio tools, and 1 model among 25 was found as the finest because it displayed optimum residues in acceptable region of RC Plot and particular quality factor (ERRAT) compared with others. Further refinement was performed inserting this finest model to GalaxyWEB program which developed 5 more 3D rectified structures. The superior model out of five 3D rectified structures was certified based on SAVES v6.0 outcome and MolProbity statistics. According to SAVES v6.0 outcome, the utmost model of the C-I-TASSER showed 86.1% residues in the most key zone of Ramachandran plot with estimating G-factors value −0.06, exhibited ERRAT score 93.08, and demonstrated 95.18% of Verify3D, respectively. Summary statistics of MolProbity demonstrated 89.76% residues in the satisfactory zone of RC plot for the best model of C-I-TASSER server (Supplementary Table A.3). Similar approach was implemented to establish the top model from the estimated 5 models of RaptorX server. RC plot for the best refined model of RaptorX server showed 87.3% of residues in most preferred area with G-factors value 0.04, exhibited overall quality factor (ERRAT) 94.90, and fewer than 80% of the amino acids have scored ≥0.2 in the 3D/1D profile. Also, MolProbity statistics displayed 95.18% residues in the preferred region of RC plot for the finest model of RaptorX server (Supplementary Table A.3). Table 3, Supplementary Table A.3 and Supplementary Fig. A.2 reveal validation statistics which indicated that RaptorX server predicted better A36R protein model of MPXV than C-I-TASSER program and can be used for further peptide-based molecular docking analysis.

**Active site analysis of A36R protein**

CASTp, DoGSiteScorer and DeepSite bio tools identified the active groove of the A36R protein of MPXV. There was detection around 12 active points of the A36R protein; those are PRO85, ASP86, THR87, ARG88, LEU90, ARG91, PHE94, ILE121, ASP122, ILE123, SER124, and TYR144 (Supplementary Fig. A.3). These amino acids of the A36R protein were meticulously evaluated to analyze molecular docking.

**Molecular docking analysis of A36R against AVPs**

All 57 models were used to dock against the A36R protein of MPXV. The docking energy from each peptide is tabulated in Supplementary Table S2. The most conducive binding scores of 10 peptide models were recorded in Table 4. Furthermore, we selected the best 3 favorable peptides among the 10 peptides for non-bonded structure analysis. The docking scores for top 3 peptide models were −6.10, −6.10 and −6.30 kcal/mol, respectively, for SATPdb10193, SATPdb21850, and SATPdb26811 peptides.

**Docking interaction between A36R and AVPs**

The non-bonded interactions between the peptides and the A36R protein of MPXV are shown in Figure 3. The SATPdb10193 peptide formed 6 hydrogen bonds with the A36R protein of MPXV at THR87, ARG91, SER124, ARG88, TYR10, and ARG88 and each of the amino acid residues had bond distance around 3.0 Å. This interaction was also stabilized by 2 electrostatic bonds at ARG88, and ASP8, 7 hydrophobic bonds at TYR54, VAL92, HIS89, PHE94, VAL92, LYS119, and ILE121, and 4 unfavorable bonds at GLU4, ASP5, ASP8, and PRO12 (Figure 3 and Table 5). The SATPdb21850 peptide and the A36R protein interaction created 6 hydrogen bonds at ASP122, SER124, GLY147, THR87, ASP120, and ILE121 showing bond distance about 2.5 Å. This complex additionally generated 1 electrostatic bond at GLU4, 2 hydrophobic bonds at ARG91 and PHE94, and 4 unfavorable bonds at TYRS, GLU4, ARG91, and SER124 (Figure 3 and Table 5). The SATPdb26811 peptide and the A36R protein interaction created 6 hydrogen bonds at ASP122, SER124, GLY147, THR87, ASP120, and ILE121 showing bond distance about 2.5 Å. This complex additionally generated 1 electrostatic bond at GLU4, 2 hydrophobic bonds at ARG91 and PHE94, and 4 unfavorable bonds at TYR5, GLU4, ARG91, and SER124 (Figure 3 and Table 5). The SATPdb26811 and the A36R protein complexation developed 10 hydrogen bonds at THR87, LYS125, ASP55, THR127, ASP122, ARG88, SER124, LYS125, and ILE123 securing approximate bond distance 3.0 Å. The SATPdb26811 and the A36R protein were also bound by 6 hydrophobic bonds at ARG91, VAL92, LEU90, LEU126, HIS89, and ILE121, and 4 unfavorable bonds at ARG88, LYS148, ASP55, and ASP122 (Figure 3 and Table 5).
Table 2. Peptide structures validation through SAVES v6.0 result and ProSA-web Z-score.

| DATABASE ID | SEQUENCE | MODEL NO. | ERRAT (%) | RC PLOT (%) | G-FACTORS | PROSA-WEB Z-SCORE |
|-------------|----------|-----------|-----------|-------------|-----------|-------------------|
| DRAMP01005  | VKSTGRADDLAVTKYLPP | Model 5   | 100       | 100         | -0.11     | -0.79             |
| DRAMP15917  | MNKIPKDLNNPG     | Model 5   | 100       | 88.9        | -0.53     | -0.64             |
| DRAMP16086  | ISFDELLDDYGESGS | Model 3   | 100       | 100         | -0.14     | -1.94             |
| DRAMP16110  | IKILGNQGSTLTGKYPSK | Model 4   | Not Applicable | 66.7    | -0.31     | -0.18             |
| DRAMP16162  | GTHPSSASGLKNDLLEN | Model 1   | 100       | 100         | -0.19     | -1.48             |
| DRAMP16553  | VVQIKOLQARIQALVYLDQ | Model 1   | 100       | 100         | 0.19      | -2.04             |
| DRAMP16556  | IKOLQARIQALVYLDQ | Model 1   | 100       | 93.8        | 0.05      | -1.86             |
| DRAMP16563  | ILAVERYLDQ      | Model 5   | 100       | 88.9        | -0.2      | -1.23             |
| DRAMP16566  | VERYLDQ        | Model 5   | Not Applicable | 83.3    | -0.35     | -1.27             |
| DRAMP16617  | MYELQKLNSWDVTNWNL | Model 3   | 100       | 100         | -0.05     | -0.75             |
| DRAMP16753  | ISOVNEKINQSLAFIRKSDELL | Model 1   | 100       | 95          | 0.1       | -1.8              |
| DRAMP16756  | VNEKINQSLAFIRKSDELL | Model 5   | 100       | 100         | 0.02      | -1.94             |
| DRAMP16798  | ITLNNS        | Model 3   | Not Applicable | 100    | 0.77      | 0.37              |
| DRAMP16880  | IEKLMKAEIRD   | Model 2   | 83.3333   | 76.2        | -0.18     | -2.61             |
| SATPdb10193 | ITFEDLLDYGP   | Model 1   | 100       | 100         | 0.12      | -2.44             |
| SATPdb10601 | GELGRLYLDDPDGYPDIHCD | Model 2   | 100       | 92.9        | -0.19     | -1.56             |
| SATPdb10735 | GFKRIRQIKDFLRNLV | Model 4   | 100       | 100         | 0.29      | -1.99             |
| SATPdb11012 | GEDLA        | Model 4   | Not Applicable | 100    | 0.74      | -0.08             |
| SATPdb13374 | INAKGVCRSTAKYVR | Model 2   | 100       | 91.7        | -0.05     | -0.19             |
| SATPdb14359 | GLVRDNMALRERLK | Model 2   | 100       | 92.3        | -0.12     | -0.98             |
| SATPdb14986 | GYSAGERI/VD  | Model 4   | 100       | 100         | 0.29      | -1.25             |
| SATPdb15048 | GQKKIRVRLS   | Model 3   | Not Applicable | 75      | 0.06      | -0.1              |
| SATPdb17545 | GIKKEKWIRQIKDFLRNLV | Model 3  | 100       | 100         | 0.25      | -1.84             |
| SATPdb17574 | GPQIREPYNWTELLL | Model 1   | 100       | 90.9        | -0.17     | -1.08             |
| SATPdb17706 | ISTTFTNTLEYPLS | Model 5   | 100       | 83.3        | -0.06     | -2.17             |
| SATPdb17733 | GELGRPYYVLDGDPYGYAT | Model 3  | 100       | 100         | -0.28     | -0.45             |
| SATPdb17990 | GIKKEKWIRQIKDFLRNLV | Model 3  | 100       | 94.4        | 0.19      | -2.24             |
| SATPdb18158 | GRFKIRKKLKLKFKKIS | Model 5   | 100       | 87.5        | 0.12      | -1.86             |
| SATPdb18474 | IRKVLFLDDIGDQADHEK | Model 1  | 100       | 94.1        | -0.05     | -1.58             |
| SATPdb18549 | VTTAQETKRGRITQKKEVI | Model 3  | 100       | 88.2        | -0.1      | 0.26              |
| SATPdb18582 | MARHRNWPLVMV  | Model 1   | 100       | 66.7        | -0.35     | -0.43             |
| SATPdb19309 | GIGVTQNYLYENQKQIANOF | Model 5  | 100       | 94.1        | 0.06      | -0.81             |
| SATPdb21810 | MDGPVQWPLTEEKIKLV | Model 3   | 100       | 93.3        | -0.11     | -0.41             |
| SATPdb21850 | ISYEY        | Model 1   | Not Applicable | 100    | -0.54     | -0.35             |
| SATPdb23879 | GMVTQYHQVLATHQEAIEKV | Model 2  | 100       | 100         | 0.13      | -1.64             |
| SATPdb24458 | VPAHKIGGNEQVDKLVSAG | Model 3  | 100       | 92.9        | -0.2      | -1.41             |

(Continued)
Virtual screening analysis of A36R against ZINC trial, ZINC fragments, and Enamine HTSC databases

Virtual screening against 3 ligand databases was used to look for specific interactions with the protein A36R (ZINC Trial, ZINC Fragments, and Enamine HTSC databases). Via virtual screening, a list of hits was formed, which was subsequently ordered by docking score (Figure 4). A lower score indicates that the ligand and the target protein have more favorable energy interactions. The top 1000 ligands with the strongest binding affinity for the therapeutic target protein A36R were reported in each drug library evaluated. The highest binding affinities were identified in ZINC Fragments (−6.9 to −7.8 kcal/mol), while the lowest binding affinities were found in the Enamine HTSC library (−8.8 to −9.8 kcal/mol) based on the docking scores. The top 100 compounds from the 3 libraries were considered for further physicochemical features and drug likeliness evaluation based on their binding affinities against A36R. The top 100 compounds had docking scores ranging from −9.3 to −9.8 kcal/mol, and all of the molecules were obtained from the Enamine HTSC database.

Pharmacokinetics (ADMET) characteristics analysis of top virtually screened molecules

Early estimation of physiologically based pharmacokinetic parameters such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) have become a significant factor in minimizing the likelihood of clinical therapeutic failure. The pharmacokinetics (ADMET) and drug likeliness qualities of all 100 medicines were examined, and the top 3 compounds for inhibiting the A36R protein of the monkeypox virus were screened. Two compounds (Z55287118 and Z97653013) exhibited outstanding Caco-2 Permeability (>−5.15 cm/s) and all 3 drugs had strong human intestine absorption (HIA) characteristics, meeting the optimum HIA value of ⩾30%, according to ADMET evaluation (Table 6). Although neither drug was able

Table 2. (Continued)

| DATABASE ID | SEQUENCE | MODEL NO. | ERRAT | RC PLOT (%) | G-FACTORS | PROSA-WEB Z-SCORE |
|-------------|----------|-----------|-------|-------------|-----------|-------------------|
| SATPdb24463 | ITTYWGLHTGERDWH | Model 5 | Not Applicable | 91.7 | −0.01 | −0.09 |
| SATPdb25248 | GELDELVYLDDPGPGPIDS | Model 3 | 100 | 92.9 | −0.23 | −1.95 |
| SATPdb25427 | VVDRGWGNGAGLFKGGSID | Model 2 | Not Applicable | 54.5 | −0.84 | 0.35 |
| SATPdb25481 | GENNELRLTRDAIE | Model 5 | 100 | 90.9 | −0.18 | −0.39 |
| SATPdb25562 | GIKEFQREFQRIKDFLRNLV | Model 4 | 100 | 100 | 0.23 | −2.22 |
| SATPdb25705 | VEGQGENNELRLTRDAIE | Model 5 | 100 | 60 | −0.32 | −0.34 |
| SATPdb25776 | IQKEIDRLNEVANKLNESLI | Model 4 | 100 | 100 | 0.1 | −2.05 |
| SATPdb26073 | IWHLQGLGHIYGETYG | Model 4 | 100 | 100 | 0.01 | −1.23 |
| SATPdb26160 | VRDQAELKLT | Model 5 | 100 | 100 | −0.3 | −1.7 |
| SATPdb26216 | GIKQFRKIVRIKDFLRNLV | Model 3 | 100 | 100 | 0.26 | −2.27 |
| SATPdb26449 | VGQPOQYYQANGGGFLI | Model 5 | 100 | 100 | −0.06 | −0.81 |
| SATPdb26452 | IKKEKVLAWVPASH | Model 4 | 100 | 75 | −0.29 | −0.48 |
| SATPdb26811 | VDGIPVSWDADARAPA | Model 2 | 100 | 90.9 | −0.37 | −0.95 |
| SATPdb27074 | IKKEKVLAWVPASHKIGN | Model 3 | 87.5 | 57.1 | −0.47 | −0.33 |
| SATPdb27355 | GSRIPTGERVWDGNGNTLLC | Model 1 | 100 | 93.3 | −0.24 | −0.92 |
| SATPdb27795 | VDGIPWEWADARAPA | Model 2 | 100 | 90.9 | −0.39 | −1.61 |
| SATPdb28063 | GSWLRIWDICEVLSDFK | Model 1 | 100 | 88.2 | −0.13 | 0.1 |
| SATPdb28187 | GWVKPAKLGG | Model 5 | Not Applicable | 100 | −0.1 | −0.83 |
| SATPdb28641 | IHVQGHNLQEVDAIGNFIPP | Model 2 | 100 | 84.6 | −0.26 | −0.86 |
| SATPdb28672 | GTKWLTEWPLTAEAECC | Model 2 | 100 | 92.9 | −0.2 | −1.25 |
| SATPdb28952 | GFKRIVQRKDFLRNLV | Model 5 | 100 | 100 | 0.29 | −1.99 |

Abbreviations: ProSA, protein structure analysis; SAVES, structural analysis and verification server; RC Plot, Ramachandran Plot.
to pass the blood-brain barrier (BBB), the 3 medicines did show considerable plasma protein binding (≤90% optimum). They were also reported to be non-inhibitors of most cytochrome P450 enzymes, non-respiratory harmful, non-carcinogenic, non-corrosive, and non-irritating to the eyes. On the other hand, only Z97653013 was shown to be non-AMES toxic. All 3 compounds were orally bioavailable (Table 6, Figure 5A), which is an important consideration in the development of oral medicines. The 3 medicines exhibit moderate clearance (2.399, 1.881, and 7.616 mL/min/kg), despite their lack of a decent T1/2 value. Despite this, the molecular weights of the compounds ranged from 483.160 to 356.120 g/mol, and the LogP values were determined to be within the ideal range of 1 to 5. The Polar surface area (TPSA) was also determined to be below the <140 Å² threshold, and all 3 agents followed the Lipinski Rule of Five.

### Table 3. Ramachandran plot statistics, ERRAT, and Verify3D scores for the A36R protein generated models using C-I-TASSER and RaptorX server.

|          | MF REGIONS① (%) | AA REGIONS② (%) | QA REGIONS③ (%) | DA REGIONS④ (%) | G-FACTORS⑤ | ERRAT SCORE⑥ (%) | VERIFY 3D SCORE⑦ (%) |
|----------|-----------------|-----------------|-----------------|-----------------|-------------|------------------|---------------------|
| C-I-TASSER | 86.1            | 9.5             | 1.3             | 3.2             | −0.06       | 93.08            | 75.60               |
| RaptorX   | 87.3            | 8.2             | 0.0             | 4.4             | 0.04        | 94.90            | Fewer than 80       |

Abbreviation: C-I-TASSER, Contact-guided Iterative Threading ASSEmblY Refinement.

①Ramachandran plot: Residues in most favored regions [A, B, L].
②Ramachandran plot: Residues in additional allowed regions [a, b, l, p].
③Ramachandran plot: Residues in generously allowed regions [−a, −b, −l, −p].
④Ramachandran plot: Residues in disallowed regions.
⑤Ramachandran plot: Overall G-factors.
⑥Overall quality factor generated by ERRAT server.
⑦Averaged 3D-1D score generated by Verify3D server.

### Table 4. The binding energy of the top 10 peptide molecules. The more negative energy signifies the favorable binding.

| DATABASE ID | SEQUENCE | BINDING SCORE (KCAL/MOL) | RMSD (Å) |
|-------------|----------|---------------------------|----------|
| SATPdb26811 | VDGIPvSWDADARAPA | −6.30                     | 0.00     |
| SATPdb10193 | ITFEDLLDYGP | −6.10                     | 0.00     |
| SATPdb21850 | ISYEY | −6.10                     | 0.00     |
| SATPdb17706 | ISTTFTNLTEYPLS | −6.00                     | 0.00     |
| SATPdb26073 | IWLHSLGQHYETYG | −6.00                     | 0.00     |
| SATPdb27795 | VDGIPvWEWDADARAPA | −6.00                     | 0.00     |
| SATPdb24463 | ITTYWGLHTGDWHL | −5.70                     | 0.00     |
| SATPdb28641 | IHVQGHLOEVDAGNFIPP | −5.70                     | 0.00     |
| DRAMP16798 | ITLNNS | −5.70                     | 0.00     |
| DRAMP16617 | MYELQKLNSWDVFNWL | −5.70                     | 0.00     |

Abbreviation: RMSD, root mean square deviation.

Molecular interactions of the selected top drug candidates with A36R protein

The docking interactions of the hit ligands (Z55287118, Z18483535, and Z97653013) with the A36R were visualized in detail with BIOVIA Discovery Studio 2021 (Table 7, Figure 5B and C). The molecules formed van der Waals interaction, conventional hydrogen bonds, and carbon hydrogen bonds with ARG91, PRO85, ARG91, SER124, ASP12, THR87, ARG88, ARG91, ASP122, and ASP120, whereas unfavorable acceptor-acceptor, pi-alkyl, and pi-anion interactions were generated with amino acids LYS125, LEU90, ARG91, ILE121, ASP118, LYS119, LYS125, PHE94, and ARG91, respectively.

Molecular dynamics simulation analysis of the top hit compounds

The top-ranking drug Z55287118 and the 3 peptides SATPdb10193, SATPdb21850, and SATPdb26811 had the best docking affinity and binding free energy; therefore, they were considered for molecular dynamics (MD) simulation to determine their binding kinetics and stability with the A36R receptors. The energy minimization, NVT, NPT equilibration methodologies were used to simulate receptor-drug complexes using the CHARMM36 force field. The RMSD (after lsq fit)
stability of the protein-drug and protein-peptide complexes demonstrated that after 2 ns of simulation, the system was consistently stabilized and inclined to remain in the plateau phase for the remainder of the time (Figure 5D, first column panel and Figure 6, first column panel). While the protein grew up quickly, the RMSD value remained steady in the range of 0.2 to 0.25 nm and 1.0 to 1.5 nm for drug and peptide, respectively. The protein-drug and protein-peptide complexes have the same RMSD pattern as the protein alone. The complexed RMSD was between 0.2 and 0.25 nm for protein-drug and 1.0 and 1.5 nm for protein-peptide. This indicates that the structures of the protein-ligand complex have already achieved a state of relative equilibrium. The complex RMSF gives each atom’s variance during the simulation (Figure 5D, second column panel and Figure 6, second column panel). The findings revealed that the binding site residues were less fluctuant for protein-drug and protein-peptide, except SATPdb10193. On average, the RMSF values were 0.4 and 0.6 nm for protein-drug and protein-peptide, respectively. A protein’s Solvent Accessible Surface Area (SASA) is a characterization of the surface area accessibility of the complex during simulation (Figure 5D, third column panel and Figure 6, third column panel). The graph illustrates how SASA for protein-drug as well as protein-peptide complexes has changed throughout time. The drug-protein complex’s SASA is lower than the proteins on its own. It means the protein is attached to the ligand. Nevertheless, for peptide-protein complex, the SASA scores were higher than the proteins. When the protein structure is exposed to solvent, it partially unfolds, as seen by the small increase in SASA values with time. SASA levels that are reduced when medications attach to receptors, on the other hand, suggest that the protein has shrunk. The radius of gyration (Rg) was estimated as a measure of structural compactness (Figure 5D, fourth column panel and Figure 6, fourth column panel). The Rg value will most likely remain consistent if the protein is folded correctly. If the protein unfolds, Rg levels will change over time. The average Rg of the protein-ligand and protein-peptide complexes was 2.0 nm, and it remained fairly stable throughout time, excluding SATPdb10193. The Rg figure shows a small decline in the total Rg value of the protein when docked with the drug and peptides, indicating that receptors in conjunction with the medicines were compactly packed and displayed stable folding. All of the data from the MD simulation of a chosen docked structure confirmed the stability of protein-ligand complex, and for protein-peptide complexes, SATPdb21850 and SATPdb26811 showed better stability.

Allergenicity and toxicity prediction

The SATPdb10193 and SATPdb26811 peptides were characterized as probable non-allergen, whereas SATPdb21850 was identified as probable allergen by AllerTOP v.2.0 bio tools. On the other hand, the SATPdb10193, SATPdb21850, and SATPdb26811 peptide molecules were non-allergic and...
| PEPTIDE NAME | AMINO ACID | INTERACTION CATEGORY | BOND DISTANCE (Å) |
|-------------|------------|----------------------|-------------------|
| SATPDb10193 | THR87      | Hydrogen bond        | 2.71469           |
|             | ARG91      | Hydrogen bond        | 2.24742           |
|             | SER124     | Hydrogen bond        | 2.79064           |
|             | ARG88      | Hydrogen bond        | 2.73909           |
|             | TYR10      | Hydrogen bond        | 3.57310           |
|             | ARG88      | Hydrogen bond        | 2.73909           |
|             |            | Electrostatic bond   | 3.10949           |
|             | ASP8       | Electrostatic bond   | 3.49650           |
|             | TYR54      | Hydrophobic bond     | 4.98725           |
|             | VAL92      | Hydrophobic bond     | 4.60781           |
|             | HIS89      | Hydrophobic bond     | 5.49742           |
|             | PHE94      | Hydrophobic bond     | 5.35305           |
|             | VAL92      | Hydrophobic bond     | 5.36524           |
|             | LYS119     | Hydrophobic bond     | 5.33564           |
|             | ILE121     | Hydrophobic bond     | 5.32334           |
|             | GLU4       | Unfavorable bond     | 0.93776           |
|             | ASP5       | Unfavorable bond     | 0.93648           |
|             | ASP8       | Unfavorable bond     | 0.93699           |
|             | PRO12      | Unfavorable bond     | 0.93667           |

| SATPdb21850 | ASP122     | Hydrogen bond        | 2.55006           |
|             | SER124     | Hydrogen bond        | 2.64414           |
|             | GLY147     | Hydrogen bond        | 2.74919           |
|             | THR87      | Hydrogen bond        | 2.44327           |
|             | ASP120     | Hydrogen bond        | 2.97542           |
|             | ILE121     | Hydrogen bond        | 2.69115           |
|             | GLU4       | Electrostatic bond   | 3.38873           |
|             | ARG91      | Hydrophobic bond     | 5.31355           |
|             | PHE94      | Hydrophobic bond     | 4.94915           |
|             | TYR5       | Unfavorable bond     | 0.93598           |
|             | GLU4       | Unfavorable bond     | 0.93633           |
|             | ARG91      | Unfavorable bond     | 4.48258           |
|             | SER124     | Unfavorable bond     | 1.84364           |

| SATPdb26811 | THR87      | Hydrogen bond        | 2.62530           |
|             | LYS125     | Hydrogen bond        | 2.00428           |
|             | ASP55      | Hydrogen bond        | 2.52361           |
|             | THR127     | Hydrogen bond        | 2.80014           |

(Continued)
non-toxic which were detected by AllergenFP v.1.0 and ToxinPred tools, respectively (Table 8).

**Discussion**

Monkeypox virus has been focused by national and international media, public and scientific attention after its recent emergence reported on May 7, 2022 in the United Kingdom,\(^{102}\) has created a pandemic situation because of its onward spread around the globe. Although different suggested drugs and vaccines are there, however, monkeypox infection is still challenging because no effective drugs or vaccines have been found that mitigate the risk of MPX infection significantly. For that, there is a necessity for rapid and effective development of active antiviral agents against MPXV. As conventional drug development procedures are time-consuming and costly, in silico–based drug design approaches were conducted using molecular docking that may promote to find out the new therapeutic agent against MPXV due to its fast and on target screening proficiency from a broad small molecule library or peptide database.\(^{103-105}\)

The A36R protein of MPXV plays significant role for the polymerization of actin tail beneath the CEV that moves the virus into the neighboring cell or the virus is freed as an EEV. Blocking the A36R protein of poxviruses inhibits the actin polymerization; consequently, viral migration, adhesion, and vesicle trafficking to the host cell are thwarted. AVPs are thought as leading component that directly target the viral proteins or inhibit the development of viruses by targeting its specific regions or components.\(^{36}\)

Our study shows that most of the selected AVPs have strong interaction with the A36R protein. It was observed that their docking scores ranging from −4.10 to −6.30 kcal/mol (Supplementary Table A.2). However, we focused on the SATPdh10193, SATPdh21850, and SATPdh26811 peptides based on docking scores for further analysis because these AVPs exhibited the higher binding affinity (−6.10, −6.10 and −6.30 kcal/mol, respectively) compared with other AVPs. The SATPdh10193 peptide had several hydrogen, electrostatic, and hydrophobic binding interactions in the catalytic residues at THR87, ARG91, SER124, ARG88, PHE94, and ILE121. These interactions at the active groove of the protein may be responsible for the higher binding energy. Researcher suggests that hydrogen, electrostatic, and hydrophobic interaction at the active site of the target protein may involve to possible inhibition of the protein.\(^{106}\) Similarly, multiple hydrogen and hydrophobic interaction in the active cavity residues at ASP122, SER124, THR87, ILE121, ARG91, and PHE94 may be responsible for the tight binding and better affinity which indicates possible blockage of the A36R protein by SATPdh21850. Moreover, the SATPdh26811 ligand showed the highest binding energy −6.30 kcal/mol while interacting with the A36R protein of MPXV. This might be result of more hydrogen and

| PEPTIDE NAME | AMINO ACID | INTERACTION CATEGORY | BOND DISTANCE (Å) |
|--------------|------------|----------------------|-------------------|
| ASP122       | Hydrogen bond | 3.29745             |
| ARG88        | Hydrogen bond | 2.66800             |
| SER124       | Hydrogen bond | 3.07170             |
| LYS125       | Hydrogen bond | 3.03728             |
| ILE123       | Hydrogen bond | 3.74970             |
| ARG91        | Hydrophobic bond | 4.29514          |
| VAL92        | Hydrophobic bond | 5.08916             |
| LEU90        | Hydrophobic bond | 4.57379             |
| LEU126       | Hydrophobic bond | 5.11871             |
| HIS89        | Hydrophobic bond | 4.69540             |
| ILE121       | Hydrophobic bond | 4.74043             |
| ARG88        | Unfavorable bond | 5.59912            |
| LYS148       | Unfavorable bond | 5.10524             |
| ASP55        | Unfavorable bond | 5.51186             |
| ASP122       | Unfavorable bond | 5.15790             |

Abbreviation: MPXV, monkeypox virus.
hydrophobic bonds with active site residues at THR87, ASP122, ARG88, SER124, ILE123, ARG91, LEU90, and ILE121.

A virtual screening evaluation of the A36R protein was also performed against the ZINC Trial library, ZINC Fragments library, and Enamine HTSC database by using Texas Advanced Computing Center's Drug Discovery portal (TACC). Despite the fact that they offered the top 1000 molecules (Figure 4) from each chemical library, we chose the top 100 based on the greatest binding affinities, with the majority of the medications from Enamine HTSC having the highest binding affinities to A36R.

The projected spike in the number of commercialized new medicines has not occurred in comparison with the number of novel drugs created each year, which has been attributed to a number of factors, including the candidate products' poor pharmacokinetic (PK) properties. As a result, effective absorption, distribution, metabolism, elimination/excretion, and toxicity (ADMET) screening criteria are required. All of the top 100 medications from the virtual screening study underwent ADMET analysis and were filtered out based on their ADMET characteristics. After screening, 3 (Z55287118, Z18483535, and Z97653013) of the 100 virtually screened compounds were eventually certified. The certified Z55287118, Z18483535, and Z97653013 molecules formed potent hydrogen, electrostatic, and hydrophobic bonds with pocket residues at ARG88, ARG91, PRO85, LEU90, ILE121, THR87, SER124, and ASP122 (Table 7). In addition, it is indisputable that Caco-2 permeability and Human Intestinal Absorption (HIA) are 2 of the most important indicators for drug absorption, and 2 of the compounds (Z55287118 and Z97653013), as well as all 3 molecules (Table 6), were shown to have satisfactory Caco-2 permeability and HIA, respectively. When the distribution properties (plasma protein binding and blood-brain barrier permeability) were investigated, almost all drugs had acceptable distribution throughout the body and conformed to the optimum values. A great oral medication should never block cytochrome p450 enzymes and be non-toxic. The majority of the compounds, on the other hand, showed good cytochrome p450 inhibition and toxicity profiles (have no Ames toxicity, non-HERG Blockers, non-respiratory toxic, and non-carcinogenic, non-irritating, and non-corrosive to the eyes). Despite this, all drugs demonstrated sufficient ADMET properties to be deemed therapeutic candidates.

As per Walters and Murcko, drug-like compounds are molecules with functional groups and physical properties that are comparable to the majority of established therapeutics, and...
Bioinformatics and Biology Insights

so may be regarded as compounds that are pharmacologically active or have therapeutic potential. According to Lipinski et al.\textsuperscript{111} the compounds are "drug-like" if they show significant acceptable ADMET properties and pass Phase I clinical studies. As a first-step screening strategy, the rule of five has been frequently used to assess the druggability of new pharmaceutical substances. Oral bioavailability is determined by the logP value and total polar surface area (TPSA), and it is important to highlight that all of the drugs were orally bioavailable (Table 6). Furthermore, all of the drugs adhered to the Lipinski rule of five as well.

In this context, we find that AVPs SATPdb10193, SATPdb21850, and SATPdb26811 as well as small molecule compounds Z55287118, Z18483535, and Z97653013 have the potentiality to block the function of the A36R protein, and such findings can be further evaluated using with additional experiments.

**Conclusion**

The study was conducted to identify the lead molecules that particularly impede the function of the A36R protein of MPXV. For that, peptide and small molecule databases were used to determine novel drug candidates against the A36R protein. Modeling, validating, docking, and finally MD simulation findings suggest the SATPdb10193, SATPdb21850, and SATPdb26811 peptide molecules have strong interaction with the active point residues of the A36R protein of MPXV. Moreover, molecular docking and simulation results of the small molecule hit compounds demonstrate that Z55287118, Z18483535, and Z97653013 had high affinities with the vital active points of the A36R protein.

**Table 6. Pharmacokinetic (ADMET) and drug likeliness evaluation of hit virtually screened drugs.**

| **ADMET PROPERTIES** | **Z55287118** | **Z18483535** | **Z97653013** | **DRUG LIKELINESS PROPERTIES** | **Z55287118** | **Z18483535** | **Z97653013** |
|----------------------|---------------|---------------|---------------|--------------------------------|---------------|---------------|---------------|
| Caco-2 permeability  | −5.246        | −4.924        | −5.286        | Molecular weight               | 356.120       | 453.080       | 483.160       |
| HIA                  | −             | −             | −             | Num. H-bond acceptors          | 5             | 8             | 10            |
| PPB                  | 100.82%       | 98.46%        | 107.05%       | Num. H-bond donors             | 2             | 1             | 0             |
| BBB penetration      | −             | −             | −             | Num. rotatable bonds           | 4             | 6             | 5             |
| CYP1A2 inhibitor     | +             | +             | −             | Number of rings                | 4             | 5             | 5             |
| CYP2C19 inhibitor    | +             | +             | −             | Number of heteroatoms          | 5             | 8             | 11            |
| CYP2C9 inhibitor     | +             | +             | +             | Num. heavy atoms               | 18            | 22            | 12            |
| CYP2D6 inhibitor     | −             | −             | −             | Formal charge                  | 0             | 0             | 0             |
| CYP3A4 inhibitor     | +             | +             | +             | Number of rigid bonds          | 25            | 32            | 32            |
| CL                   | 2.399         | 1.881         | 4.466         | Flexibility                    | 0.160         | 0.188         | 0.156         |
| T<sub>1/2</sub>      | 0.068         | 0.177         | 0.476         | Number of stereocenters        | 0             | 0             | 0             |
| Human, oral bioavailability | 0.7857 | 0.5571 | 0.5857 | TPSA                          | 75.270        | 119.750       | 125.480       |
| Hepatotoxicity       | −             | −             | −             | logS                          | −7.404        | −6.373        | −4.364        |
| Respiratory toxicity | −             | −             | −             | logP                          | 5.082         | 4.195         | 1.355         |
| AMES toxicity        | +             | +             | −             | Lipinski rule                  | Accepted      | Accepted      | Accepted      |
| Skin sensitization   | −             | −             | −             | Pfizer rule                   | Accepted      | Accepted      | Accepted      |
| Carcinogenicity      | −             | −             | +             | GSK rule                      | Rejected      | Rejected      | Rejected      |
| Eye corrosion        | −             | −             | −             | Golden Triangle                | Accepted      | Accepted      | Accepted      |
| Eye irritation        | −             | −             | −             | PAINS                         | 1 alert       | 1 alert       | 0 alert       |

Abbreviations: ADMET, absorption, distribution, metabolism, excretion, and toxicity; BBB, blood-brain barrier; CL, clearance rate; HIA, human intestine absorption; PPB, plasma protein binding; TPSA, total polar surface area.
Figure 5. Bioavailability radar, binding modes, molecular interaction, and Trajectory plots of MD simulation of the top hit compounds: (A) Oral-bioavailability profile of 3 hit ligands, where the yellow area indicates the upper limit, pink area denotes lower limit, and the blue lines signifies our compound properties; (B) hit ligands’ binding modes with receptor; (C) interactions of the hit compounds with receptor amino acid residues. The color bars represent the type of interaction formed between ligand and the receptor. (D) Plot of root mean square deviation (RMSD) (first column panel), root mean square fluctuation (RMSF) (second column panel), solvent accessible surface area (SASA) with respect to time (picoseconds) (third column panel), and radius of gyration (Rg) (fourth column panel) during MD simulation of A36R-receptor in complex with the hit ligand Z55287118(red color) and unbound protein (green color). MD indicates molecular dynamic.

Table 7. Molecular interactions of the hit drugs with receptor A36R.

| DRUG CANDIDATES | AMINO ACIDS INVOLVED IN H BONDING | ELECTROSTATIC/HYDROPHOBIC INTERACTIONS INVOLVED |
|-----------------|----------------------------------|-----------------------------------------------|
| Z55287118       | ARG91, PRO85                     | LYS125, LEU90, ARG91, ILE121                  |
| Z18483535       | THR87, ARG91, SER124, PRO85, ASP122 | ASP118, ILE121, ARG91, LYS125                 |
| Z97653013       | THR87, ARG88, ARG91, ASP122, ASP120 | LYS119, LYS125, PHE94, ILE121, ARG91           |
Figure 6. Trajectory plots of MD simulations: Plot of root mean square deviation (RMSD) (first column panel), root mean square fluctuation (RMSF) (second column panel), solvent accessible surface area (SASA) with respect to time (picoseconds) (third column panel), radius of gyration (Rg) (fourth column panel), and hydrogen bond (fifth column panel) during MD simulation of A36R in complex with peptides SATPdb10193, SATPdb21850, SATPdb26811 (red color), or unbound protein A36R (blue color). MD indicates molecular dynamic.
However, these assessments, additionally future in vivo and in vitro clinical trials, can lead to the coherent and rigorous development of the peptide and molecular based inhibitors by targeting the A36R protein of MPXV.

Acknowledgements
We acknowledge high performance computing facility support from Centre for Bioinformatics Learning Advancement and Systematics Training (cBLAST) and Texas Advanced Computing Center (TACC) for MD analysis. We also acknowledge support of Biomolecular Research Foundation (BMRF), Dhaka, Bangladesh.

Author Contributions
Abul Bashar Mir Md. Khademul Islam: Conceptualization, Methodology, Software, Writing- Original draft preparation, Visualization, Investigation, Supervision, Validation, Writing-Reviewing and Editing. Mohammad Mohasin Miah: Data curation, Methodology, Software, Writing- Original draft preparation, Visualization, Validation, Writing- Reviewing and Editing. Nuzhat Tabassum: Data curation, Methodology, Software, Writing- Original draft preparation, Visualization, Validation, Writing- Reviewing and Editing. Malika Afroj Zinnia: Data curation, Methodology, Software, Writing- Original draft preparation, Visualization, Investigation, Validation, Writing- Reviewing and Editing.

Data and Software Availability
All data added: Tables, figures, and supplementary file and supplementary tables. In this research work, publicly available free mostly online and few offline software/tools were used. Necessary link and reference of the software/tools are provided in the method section.

Supplemental Material
Supplemental material for this article is available online.

REFERENCES
1. Ladnyj JD, Ziegler P, Kima E. A human infection caused by monkeypox virus in Basankusu Territory, Democratic Republic of the Congo. Bull World Health Organ. 1972;46:593-597.
2. Reed KD, Melski JW, Graham MB, et al. The detection of monkeypox in humans in the Western Hemisphere. N Engl J Med. 2004;350:342-350.
3. Mazurkov OY, Kabanov AS, Shishkina LN, et al. New effective chemically synthesized anti-smallpox compound NIOCH-14. J Gen Virol. 2016;97:1229-1239.
4. Dzurki KN, McCollum AM, Nakazawa Y, et al. Emergence of monkeypox—West and Central Africa, 1970–2017. Med Microbiol Med Hyg. 2018;67:306-310.
5. Parker S, Nuara A, Buller RM, Schultz DA. Human monkeypox: an emerging zoonotic disease. Future Microbiol. 2007;2:17-34.
6. Arita I, Henderson DA. Smallpox and monkeypox in non-human primates. Bull World Health Organ. 1986;63:277-283.
7. Likos AM, Sammons SA, Olson VA, et al. A tale of two clades: monkeypox viruses. J Gen Virol. 2005;86:2661-2672.
8. Reynolds MG, Damon IK. Outbreaks of monkeypox after cessation of smallpox vaccination. Trends Microbiol. 2012;20:80-87.
9. Vidal-Giralt DB, Eckburg PB. Human monkeypox: an emerging zoonosis. Lancet Infect Dis. 2004;4:145-146.
10. Nolen LD, Osadahe L, Kotomba J, et al. Extended human-to-human transmission during a monkeypox outbreak in the Democratic Republic of the Congo. Emerg Infect Dis. 2016;22:1014-1021.
11. Rimston AW, Mulembakani PM, Johnston SC, et al. Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic Republic of Congo. Proc Natl Acad Sci USA. 2010;107:16262-16267.
12. Sale TA, Melski JW, Stratman EJ. Monkeypox: an epidemiologic and clinical comparison of African and US disease. J Am Acad Dermatol. 2006;55:478-481.
13. McCollum AM, Damon IK. Human monkeypox. Clin Infect Dis. 2014;58:260-267.
14. Breman JG, Steniowski MV, Zanotto E, Gromyko AI, Ariza I. Human monkeypox, 1970-79. Bull World Health Organ. 1980;58:165-182.
15. Yinka-Ogunleye A, Aruna O, Dalhat M, et al. Outbreak of human monkeypox in Nigeria in 2014: a clinical and epidemiological report. Lancet Infect Dis. 2019;19:872-879.
16. Guarnier J, Johnson BJ, Paddock CD, et al. Monkeypox transmission and pathogenesis in prairie dogs. Emerg Infect Dis. 2004;10:426-431.
17. Arita I, Jezeck Z, Khodakevich L, Ruti K. Human monkeypox: a newly emerged orthopoxvirus zoonosis in the tropical rain forests of Africa. Am J Trop Med Hyg. 1985;34:781-789.
18. Breman JG, Henderson DA. Diagnosis and management of smallpox. N Engl J Med. 2002;346:1300-1308.
19. Stagles KL, Watson AA, Boyd JF, More IA, McSeveney D. The histopathology and electron microscopy of a human monkeypox lesion. Trans R Soc Trop Med Hyg. 1985;79:192-202.
20. WHO. The current status of human monkeypox: memorandum from a WHO meeting. Bull World Health Organ. 1984;62:703-713.
21. Shchelkunov SN, Totmenin AV, Safronov PF, et al. Analysis of the monkeypox virus genome. Virology. 2002;297:172-194.
22. Buller RM, Palumbo GJ. Poxvirus pathogenesis. Microbiol Rev. 1991;55:80-122.
23. Roberts KL, Smith GL. Vaccinia virus morphogenesis and dissemination. Cell Host Microbe. 2008;13:472-479.
24. Rottner K, Stradal TEB. Poxviruses taking a ride on actin: new users of known motors. Trends Microbiol. 2009;16:472-479.
25. Fields BN, Knipe DM, Howley PM. Fields Virology. 2013. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.
26. Moss B. Membrane fusion during poxvirus entry. Cell Host Microbe. 2009;6:497-499.
27. Fields BN, Knipe DM, Howley PM. Fields Virology. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.
28. Donnelly SK, Weisowage I, Zettl M, Way M. WIP provides an essential link between Nck and N-WASP during Arp2/3-dependent actin polymerization. Curr Biol. 2013;23:999-1006.
29. Fosgerau K, Hoffmann T. Poxvirus therapeutics: current status and future directions. Drug Discov Today. 2015;20:122-128.

Table 8. The allergic and toxicity profile of the top three peptide molecules.

| PEPTIDE ID | ALLERGENICITY PREDICTION | TOXICITY PREDICTION |
|-----------|--------------------------|---------------------|
| SATPdb10193 | Probable non-allergen | Non-toxin |
| SATPdb21850 | Probable allergen | Non-toxin |
| SATPdb28811 | Probable allergen | Non-toxin |
94. Bondi A. Van der Waals volumes and radii. J Phys Chem. 1964;68:441-451.
95. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9:90-95.
96. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585:357-362.
97. R Core Team. A language and environment for statistical computing. Vienna: R Project for Statistical Computing; 2020.
98. Osorio D, Rondon-Villacarril P, Torres R. Peptides: a package for data mining of antimicrobial peptides. R J. 2015;7:4-14.
99. CDC. Human monkeypox—Kasai oriental, democratic Republic of Congo, February 1996-October 1997. MMWR Morb Mortal Wkly Rep. 1997;46:1168-1171.
100. Schoch CL, Ciufo S, Domrachev M, et al. NCBI taxonomy: a comprehensive update on curation, resources and tools. Database. 2020;2020:baa062.
101. van de Waterbeemd H, Gifford E. ADMET in silico modelling: towards prediction paradise? Nat Rev Drug Discov. 2003;2:192-204.
102. WHO. Multi-country monkeypox outbreak in non-endemic countries. https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON385. Published 2022. Accessed May 21, 2022.
103. Kalhor H, Sadeghi S, Abolhasani H, Kalhor R, Rahimi H. Repurposing of the approved small molecule drugs in order to inhibit SARS-CoV-2 S protein and human ACE2 interaction through virtual screening approaches. J Biomol Struct Dyn. 2022;40:1299-1315.
104. Mathpal S, Joshi T, Sharma P, et al. A dynamic simulation study of FDA drug from zinc database against COVID-19 main protease receptor. J Biomol Struct Dyn. 2022;40:1084-1100.
105. Khan MA, Mahmud S, Ul Alam ASMR, et al. Comparative molecular investigation of the potential inhibitors against SARS-CoV-2main protease: a molecular docking study. J Biomol Struct Dyn. 2020;39:1-7.
106. Hasan MA, Mazumder MH, Chowdhury AS, Datta A, Khan MA. Molecular docking study of malaria drug target enzyme transketolase in Plasmodium falciparum 3D7 portends the novel approach to its treatment. Source Code Biol Med. 2015;10:7.
107. Kennedy T. Managing the drug discovery/development interface. Drug Discov Today. 1997;2:436-444.
108. Vardhan S, Sahoo SK. In silico ADMET and molecular docking study on searching potential inhibitors from limonoids and triterpenoids for COVID-19. Comput Biol Med. 2020;124:103936.
109. Norinder U, Bergström CAS. Prediction of ADMET properties. ChemMedChem. 2006;1:920-937.
110. Walters WP, Murcko MA. Recognizing molecules with drug-like properties. Curr Opin Chem Biol. 1999;3:384-387.
111. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46:3-26.