Discovery of a Heparan-Sulfate Binding Domain in Monkeypox Virus H3 Protein as a New Anti-poxviral Drug Target

Bin Zheng[1], Meimei Duan[2], Zhuojian Lu[1], Lichao Liu[1], Shangchen Wang[1], Yifen Huang[1], Guojin Tang[1], Lin Cheng[2]* and Peng Zheng[1]*

1. State Key Laboratory of Coordination Chemistry, Chemistry and Biomedicine Innovation Center (ChemBIC), School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, Jiangsu, 210023, China
2. Institute for Hepatology, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Shenzhen, 518112, China

*Email: pengz@nju.edu.cn & chenglin252@163.com

Abstract: The 2022 global outbreak of mpox virus (MPXV), recognized as a public health emergency by the World Health Organization (WHO), underscored the urgent need for specific antiviral agents. H3, a key antiviral target, mediates the adhesion of MPXV to host cells through its interaction with heparan sulfate (HS). Understanding the details of the H3-HS interaction has been challenging due to the highly flexible and complex nature of HS, which hinders its structural characterization. Utilizing state-of-the-art computational biology techniques, including molecular dynamics (MD) simulations, dynamic molecular docking, and AlphaFold2, we identified a novel α-helical domain in H3 critical for HS binding. This domain, conserved across *orthopoxviruses*, was validated experimentally, confirming its crucial role in viral adhesion at both molecular and cellular levels. Targeting this newly discovered HS-binding domain, we *de novo* designed a protein inhibitor, AI-PoxBlock723, which effectively disrupt the H3-HS interaction, and inhibit viral infection. Our work not only introduces a promising anti-poxviral drug candidate, but also provides a comprehensive framework for elucidating complex protein-glycan interactions.

Introduction

Mpox virus (also known as monkeypox virus), a member of the *Orthopoxvirus* genus within the *Poxviridae* family, is an emerging zoonotic pathogen. This genus also includes the Variola virus (VARV), the causative agent of smallpox, and Vaccinia virus (VACV), which has been used as a live vaccine against smallpox[1,2]. Historically endemic to Central and West Africa, MPXV gained global attention in 2022 when it spread internationally, prompting the World Health Organization (WHO) to declare it a public health emergency of international concern[3,4]. Unlike smallpox, MPXV has a lower fatality rate of 2-10%, yet it continues to pose significant health challenges due to the absence of specific antiviral treatments[5-10].

The adhesion of MPXV to host cells, mediated through the viral adhesion protein, represents the critical first step in viral infection[11-13]. For example, H3 binds to cell-surface heparan sulfate (HS), which is not only pivotal for MPXV but also represents a potential universal target across *orthopoxviruses*. Therefore, H3-specific neutralizing monoclonal antibody protects rabbits and immunodeficient mice against lethal poxvirus infections[14], and novel mpox vaccines have been developed recently with H3 as an important component[15,16]. However, the structural intricacies of the H3-HS interaction are not well-understood, largely due to the dynamic and heterogeneous nature of HS. This has significantly hindered the development of targeted therapeutics.

Recent advances in computational biology have opened new avenues for elucidating complex biological interactions. This study employs state-of-the-art AI-enhanced structural predictions, all-atom molecular dynamics simulations, and dynamic molecular docking to uncover novel insights into the H3-HS interaction. We identified here a previously uncharacterized α-helical domain within H3 which is essential for effective HS binding[17-21]. Further biophysical and biochemical validation through single-molecule force spectroscopy, flow cytometry and biolayer interferometry has confirmed the functional importance of this domain in the viral adhesion process[22-25]. Based on this discovery, we developed a de novo designed protein inhibitor targeting the α-helical domain of H3, demonstrating efficacy in viral inhibition assays[26,27]. These findings not only highlight a promising therapeutic approach against MPXV and potentially other pathogens utilizing similar adhesion mechanisms but also demonstrate the powerful capabilities of computational biology, particularly AI-assisted methodologies, in the discovery of specific antiviral agents..

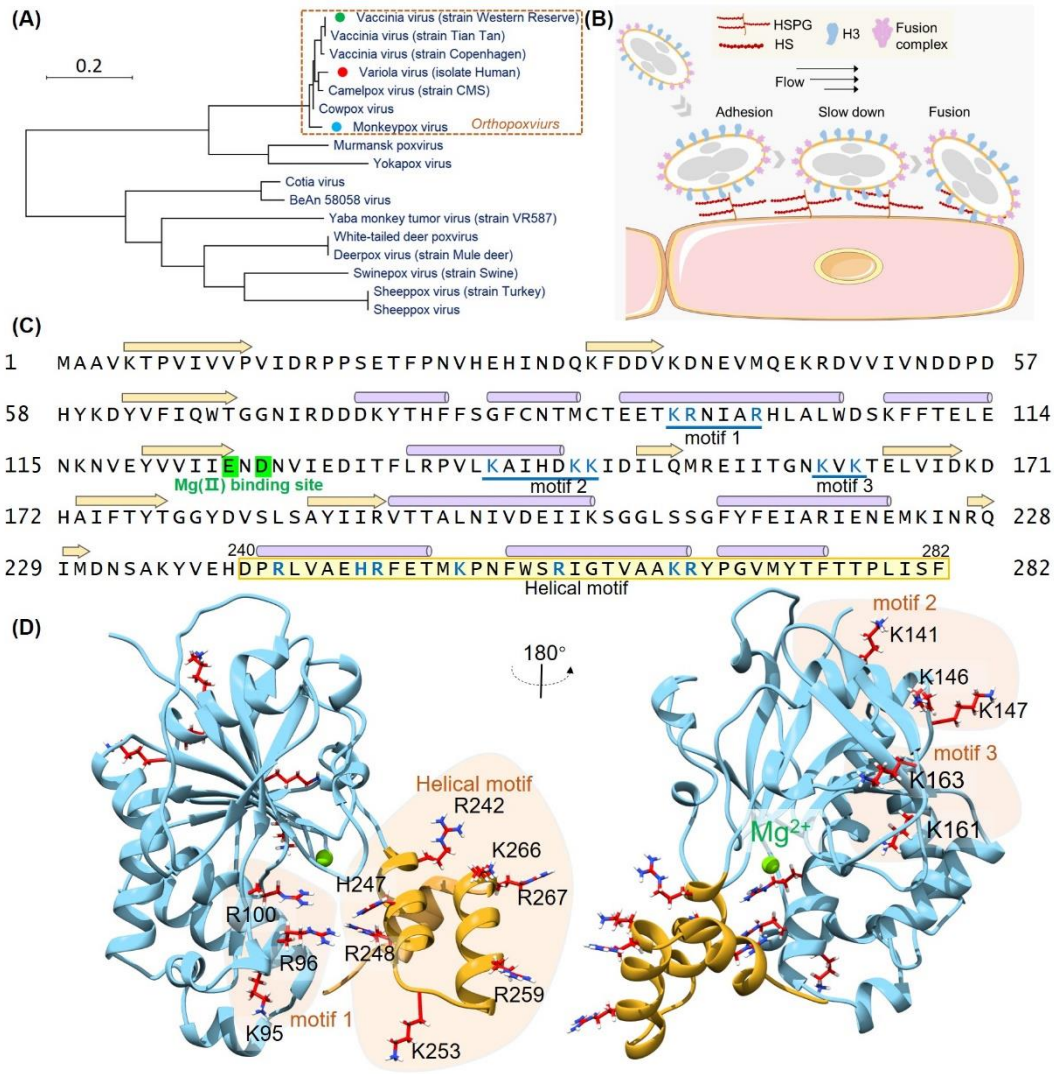


Fig. 1 **Structural Analysis of the H3 Protein in Monkeypox Virus** (A) Phylogenetic tree depicts the evolutionary relationships within the *Poxviridae* family, highlighting MPXV, (blue circle), VARV, (red circle), and VACV, (green circle). (B) Schematic shows MPXV adhesion to a cell, with H3 and fusion complexes depicted. (C) The amino acid sequence and structure of MPXV H3 displays β-strands (yellow arrows) and α-helices (pink cylinders). Mg(II) binding sites (green) and potential HS binding motifs (blue underlines) are shown. The novel three-helical domain (240-282), is highlighted in yellow. (D) AlphaFold2 prediction of MPXV H3 Structure shownin two orientations. Regions matching VACV H3's known structure are in blue; novel elements predicted are in yellow, with all potential GAG-binding motifs highlighted.

Result

Identification of the Novel HS Binding Site through Comprehensive Computational Biology Methods

The H3 protein of MPXV is a 324-amino-acid transmembrane protein and conserved across all *orthopoxviruses*. Pioneer experiments have shown that its extracellular domain (residues 1-284) binds to cell-surface heparan sulfate (HS)[28,29],

facilitating the viral particle's entry in conjunction with the fusion protein system[30]. Therefore, inhibiting HS binding to H3 represents a promising strategy for anti-poxviral drug development. Although the structure of the extracellular region of VACV H3 is known, the dynamic and heterogeneous nature of HS complicates the identification of precise binding sites and mechanisms, challenging the development of targeted antiviral treatments.

To elucidate potential HS binding sites, we initially analyzed the sequence and structure of H3. The AI-based AlphaFold2 has proven to be a powerful tool for protein structure prediction and is increasingly used for protein-target drug development. Given that only the H3 structure from VACV was available[28], we utilized AlphaFold2 to model the MPXV H3 structure (Fig. S1), focusing on regions rich in basic amino acids—K (Lysine), R (Arginine), and H (Histidine)—which are known to interact with glycosaminoglycans (GAG) like HS. We identified three potential HS binding motifs first: Motif 1 (K95, R96, R100), Motif 2 (K141, K146, K147), and Motif 3 (K161, K163)[31], each located within the known homologous structure (VACV H3) (Fig. 1C, D).

Interestingly, a cluster of seven additional positively charged residues (R242, H247, R248, K253, R259, K266, R267) in the C-terminal region was revealed. This region, not structurally characterized previously due to its removal in X-ray diffraction studies, was retained in all other biochemical and viral assays[28-30]. Alphafold2 predicted these residues to form a three-helical domain, which may represent a novel binding site for HS (Fig. 1C, D).

To validate these predicted binding sites, we employed flexible molecular docking followed by extensive molecular dynamics (MD) simulations[32,33]. We selected a 20-repeat HS unit with a specific composition for detailed analysis ($[\text{IdoA2S-GlcNS6S-IdoA-GlcNS(3,6S)}]_5$) (Fig. S2). A 500 ns MD simulation established the equilibrium conformation of H3, and subsequent docking with AutoDock Vina identified four high-scoring conformations for each motif (Fig. 2A, Fig. S3). To address the dynamic nature of HS, we extended our simulations to 3 times additional 1000 ns MD simulations for each conformation. RMSD (Root Mean Square Deviation) analysis demonstrated notably stable binding for HS to both Motif 1 and the newly identified helical domain (Supplementary Video 1), with RMSD values consistently below 4 nm, indicating minimal structural fluctuations during interaction (Fig. 2B).

Moreover, we quantitatively analyzed the binding free energy between H3 and HS for different conformational states using umbrella sampling techniques. A harmonic potential was applied to HS and gradually pulled away from H3, generating a series of reaction coordinates (Fig. 2C, Fig. S4, Supplementary Video 2). The green force-extension curve represents the force during this stretching process. This curve, along with the corresponding PMF (Potential of Mean Force) variation shown in Figure S4, demonstrated the free energy changes throughout the umbrella sampling process, revealing that the helical domain possesses the highest binding affinity with a ΔG of -45 kCal/mol (Fig. 2D).

Although the exact binding site between H3-HS is unknown, previous studies have suggested the critical role of the Mg(II) ion in stabilizing HS binding, where its removal reduces binding efficiency[28]. Indeed, our newly discovered α-helical domain is very close to the Mg(II)-bound region in H3. Notably, these two parts appear to form a distinct cavity-like binding pocket for HS (Fig. 1D, Fig. 2E). Thus, we conducted a 1000 ns REMD (Replica Exchange Molecular Dynamics) simulation of the H3-HS complex to further understand the binding dynamics. REMD, an advanced sampling technique, helps overcome the kinetic barriers typically encountered in conformational transitions by utilizing a series of temperature-controlled replicas to enhance the exploration of the conformational landscape[34]. For this study, we set up 64 replicas over a temperature range from 310K to 387K, cumulatively simulating for 64 μs. This comprehensive simulation captured the HS molecule progressively binding deeper into a cavity formed by the α-helical domain, eventually interacting with the Mg(II) ion as hypothesized (Fig. 2E, Supplementary Video 3).

The free energy landscape, mapped out from these simulations, shows HS transitioning from a stable interaction with the helical domain across an energy barrier into the Mg(II)-enhanced cavity. This dual interaction—first with the helical domain and then with the Mg(II) site—was consistently observed to stabilize the complex further, evidenced by a significantly lower binding free energy of -67 kCal/mol (Fig. S4)[35]. These results underscore the dynamic nature of the H3-HS interaction and validate our model of sequential binding, which could be critical for designing inhibitors that target these specific interactions.

Given the crucial role of electrostatic interactions in the HS-H3 binding process, we conducted an exhaustive analysis of salt bridge formations across all MD simulation trajectories. This analysis focused on the formation of salt bridges over time between the basic amino acids (R, K, H) of H3 and HS, with brighter areas indicating more frequent formation of salt bridges and robust interaction (Fig. 2F, Fig. S5). A bar graph accompanied these maps, quantifying the average number of salt bridges each amino acid formed during the simulations, further illustrating the intensity and frequency of these interactions. Our findings demonstrated a pronounced affinity of HS for the helical domain of H3, with basic amino acids in this region forming significantly more salt bridges compared to other parts of the protein. The surface hotspot map of H3 (Fig. 2G) visually highlighted these interactions, with regions of frequent salt bridge formation represented in varying shades of blue, underscoring their distribution and intensity. In the helical domain, a focus view showed that the 1st, 7th, 9th, and 18th sugar units of HS formed salt bridges with R248, R242, R267, and R259 of H3, respectively (Fig. 2H, left), while the sulfate group on the 2nd sugar side chain of HS bound to Mg(II) (Fig. 2H, right).

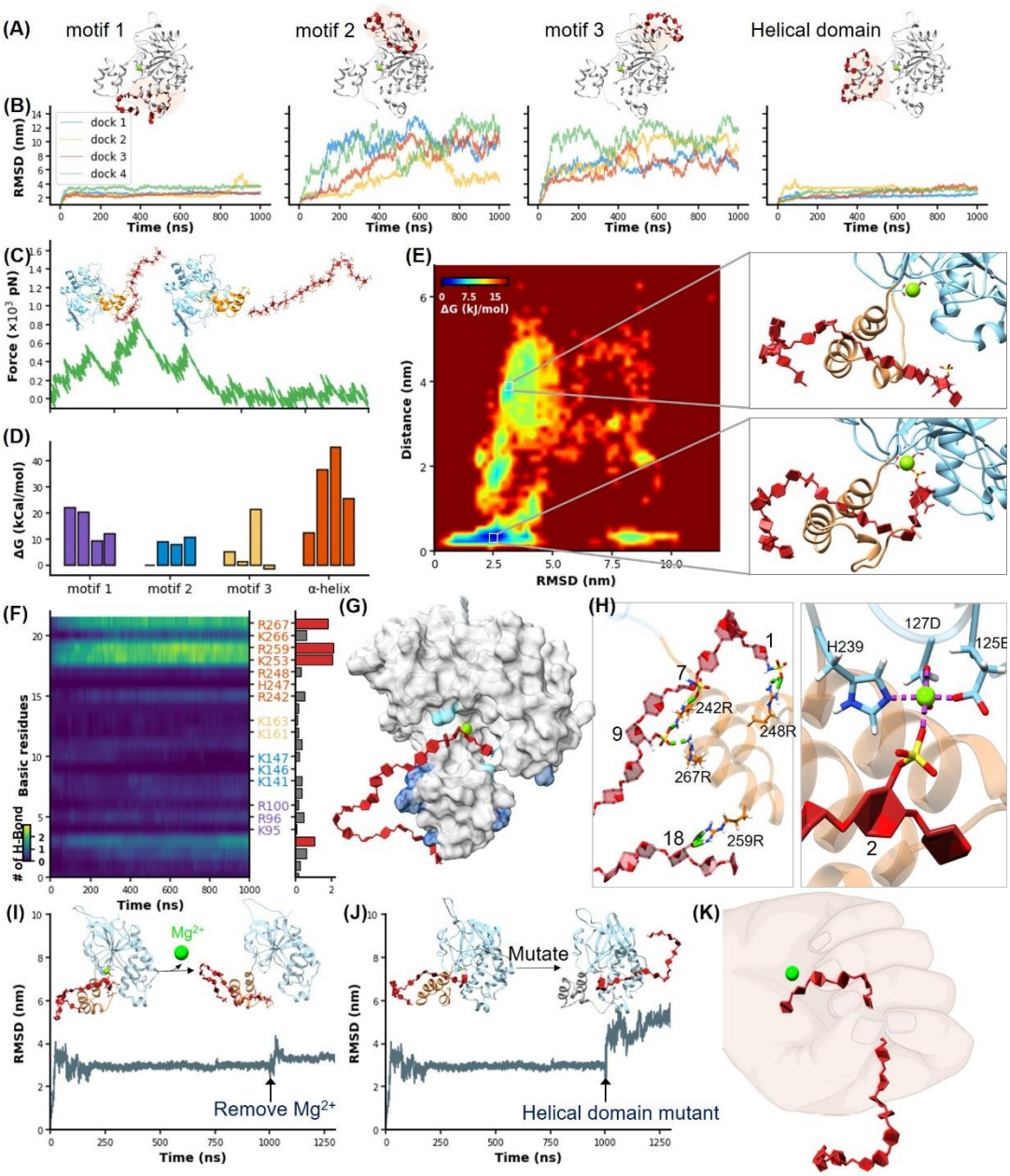


Figure 2: Molecular Dynamics and Docking Analyses of H3-HS Interactions (A-B) Cartoons show docking results of heparan sulfate (HS) with H3 motifs 1, 2, 3, and the helical domain, respectively. Panels (B) show RMSD values from 1μs MD simulations, color-coded to match the configurations in (A). (C) Schematics illustrates the reaction coordinate in umbrella sampling, highlighting HS dissociation from H3 with a green force curve. (D) Presents binding free energies for HS-H3 interactions in motifs 1, 2, 3, and the helical region, color-coded respectively. (E) A free energy landscape map from a 1000 ns REMD simulation shows HS binding configurations to the helical and Mg(II) regions. (F) Panel provides salt bridge formation statistics between HS and H3's basic amino acids, with a bar chart of average formations. (G) A surface plot shows the frequency of salt bridge formations within H3, with areas of frequent formations in blue. (H) Detailed views of HS-H3 interactions, with the left image showing salt bridges and the right image displaying electrostatic interactions

with Mg(II). This panel illustrates the impact on HS binding stability to H3 following the removal of Mg(II) during the simulation. (J) The effects of mutating all basic amino acids in the helical domain on the binding stability of HS are shown. (K) The "palm-binding" model is depicted where HS is secured by the helical "fingers" and interacts with the Mg(II) "palm."

To further substantiate these findings, we performed two control experiments. First, we removed the Mg(II) for MD simulation. Based on the results from umbrella sampling, we extended the simulation of the most stable H3-HS binding conformation to 1000 ns, then removed the Mg(II) ion and continued the simulations. The resulting increase in fluctuations of the RMSD values confirmed the stabilizing role of Mg(II) in the binding process (Fig. 2I, Supplementary Video 4). Second, we mutated all seven basic amino acids in the helical domain to serine and observed significant alterations in HS binding dynamics, supporting the critical role of these positively charged residues in stabilizing the interaction (Fig. 2J, Supplementary Video 5).

Thus, the interaction model revealed that the α-helical domain of H3 functions like a thumb, guiding HS into a stable binding position, while the Mg(II) region acts like a palm, securing the HS in place. This dynamic interaction allows HS to transition from initially binding at the protein surface to being deeply anchored within the protein cavity, illustrating a complex but well-organized binding process (Fig. 2K).

Analysis of the H3 protein sequence across the *Poxviridae* family highlights its conservation and significance in HS binding. Using NCBI's BLAST, we retrieved sequences of H3 from 66 *Poxviridae* members and performed multiple sequence comparisons[36]. Structural predictions via AlphaFold2 demonstrated that these sequences generally share a similar architecture to MPXV H3, particularly the presence of the α-helical domain (Fig. S6). Given the crucial role of electrostatic interactions in the H3-HS binding process, we further analyzed the charge distribution within this domain. A heatmap detailing the side-chain charge distribution at a physiological pH of 6.5—mimicking the microenvironment of H3-HS interaction—revealed a significant concentration of positive charges at the C-terminal end of the helical domain (Fig. 3A).

Notably, this domain consistently exhibited an accumulation of basic amino acids such as lysine, arginine, and histidine, which are essential for binding the negatively charged HS. These amino acids, including residues identified as 358, 361, 367, 373, 379, 386, and 387 in *Poxviridae* H3 (global position in alignment), corresponding to residues H239, R242, R248, K253, R259, K266, R267 in MPXV H3, show high conservation (Fig. 3B, S7). Surface charge analysis of H3 proteins from four *orthopoxviruses*, including MPXV, VACV, VARV and CPXV confirmed the predominance of strong positive charges in their α-helical domains (Fig. 3C), crucial for effective binding to HS. This consistent feature across different viruses underscores the evolutionary importance of the helical domain, suggesting a universal mechanism in orthopoxviral adhesion that could be targeted in antiviral strategy.

Experimental Confirmation of the Structure and Function of α-Helical Domain

To validate our computational findings, we first conducted experiments using atomic force microscopy-based single-molecule force spectroscopy (AFM-SMFS). This technique, widely used for studying protein (un)folding, allowed us to directly assess the structure and stability of the α-helical domain within the H3 protein[37-39]. We engineered the H3 construct with two GB1 domains, each 18 nm in length upon unfolding, and one Cohesin module for precise measurement (Fig. S8)[40-42]. During the experiment, the coated AFM tip with GB1 and Dockerin initiated a Coh-Doc interaction, producing a characteristic sawtooth-like force-extension curve as it was stretched (Fig. 3D, 3E). Notably, two peaks corresponding to the unfolding of the α-helical domain were observed. The first peak showed a contour length increment (ΔLc) of 13 nm, closely matching the theoretical unfolding length for the 42 amino acid-long α-helical domain of H3 (residue 282-240, 42*0.36 nm/aa-2.6 nm)[43]. The measured unfolding force was 67.4 ± 2.1 pN (n=139), while the second peak indicated the unfolding of the remaining structure of H3, with a ΔLc of 84 nm (Fig. 3F). The total unfolding, represented by a cumulative ΔLc of 97 nm, confirmed the structured and stable nature of the α-helical domain within H3.

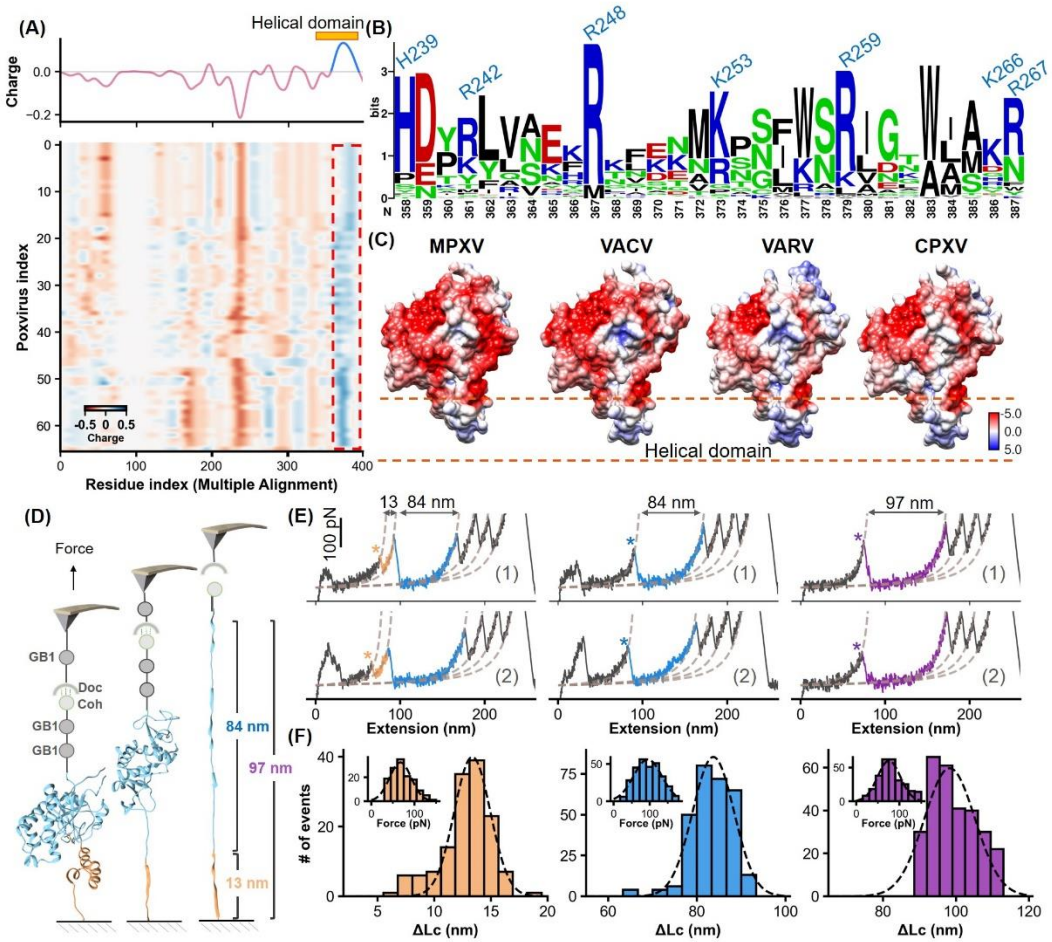


Fig. 3 **Charge Characteristics and Structural Analysis of H3 Protein** (A) The heatmap shows the amino acid charge distribution in the H3 protein across 66 *Poxviridae* viruses, following multiple sequence alignment. Blue indicates areas with more positive charges, and red indicates more negative charges. The accompanying

curve shows the average charge of all amino acids in *Poxviridae* H3. (B) Logo plot of the amino acid sequence of the helical region, highlighting the conservation of basic amino acids at specific positions. (C) The surface charge analysis of H3 from the MPXV, VACV, VARV, and cowpox viruses (CPXV) with the helical domain showing a significantly positive charge. (D) Schematic of the single-molecule force spectroscopy unfolding experiment on H3, illustrating the unfolding process of the helical region (yellow) followed by the main body (blue). (E) Representative curves of H3 unfolding, color-coded to show the helical region (yellow), main body (blue), and full-length (purple) unfolding. (F) Histograms depict the force spectroscopy signals for helical domain, main body, and full-length unfolding, with ΔLc statistics provided. The inset shows a Gaussian fit of unfolding forces.

We further explored the functional role of the α-helical domain in mediating H3's binding to HS using AFM on live cells[44,45]. The H3 protein was linked to the AFM tip and approached cultured CHO-K1 cells, which express HS, mounted on a Petri dish. Using inverted microscopy for precise positioning, the interaction between H3 and HS was initiated upon contact, subsequently monitored by retracting the tip to generate a force-extension curve. This process highlighted a distinct peak indicative of H3-HS binding with a dissociation force of 33.7 ± 0.2 pN and a binding probability of 43% (Fig. 4A, B). Trials performed across different cell surface areas produced a detailed force map, quantifying the distribution of HS unbinding forces and confirming the essential role of the helical domain in viral adhesion (Fig. 4C, D).

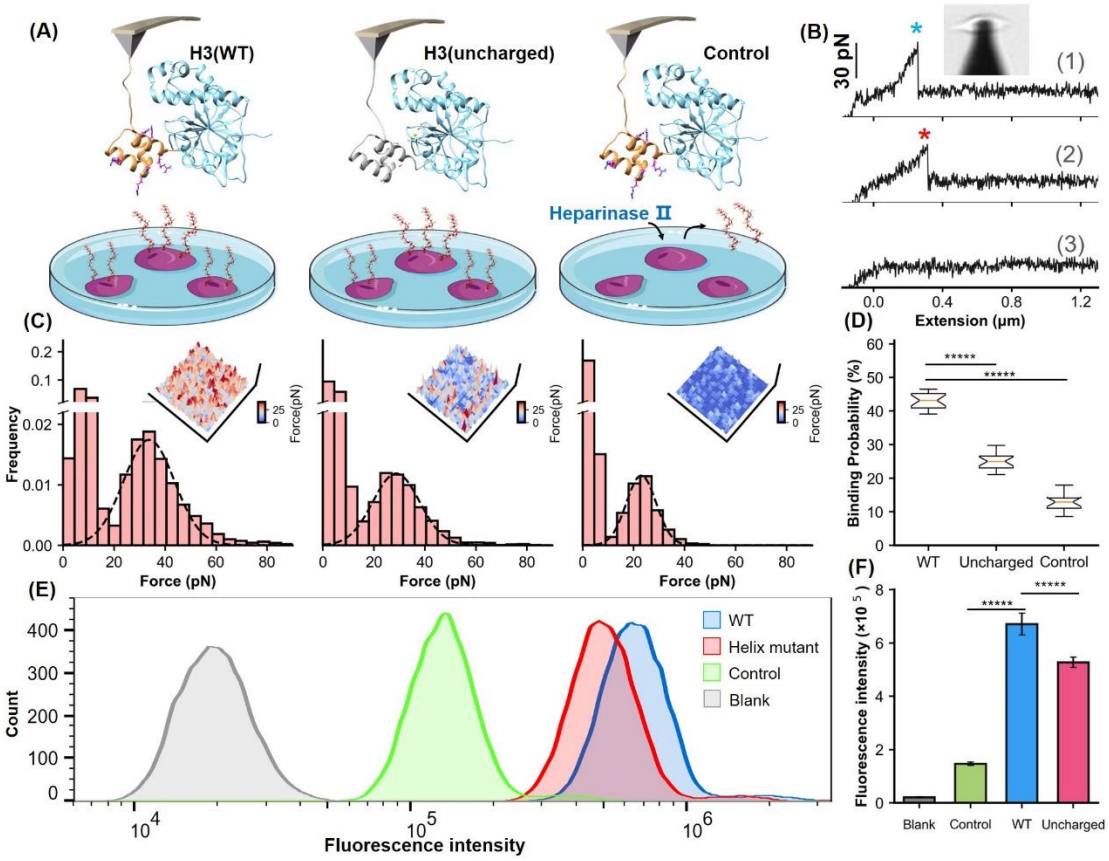


Fig. 4 **Analysis of the Helical Domain's Interaction with HS at Cellular Level** (A) Schematic of the cell force spectroscopy experiment setup shows three scenarios: wild-type H3 on an AFM tip interacting with HS on CHO-K1 cells, mutation of all basic amino acids in the H3 helical region to serine, and cells treated with HS hydrolase to remove surface HS before testing. (B) Force-extension curves depict interactions for the wild type, mutant, and control groups, marked with blue and red asterisks for dissociation events. An inset shows the optical microscope positioning the AFM probe. (C) Histograms of dissociation signals comparing the wild type, mutant, and control groups, with an inset detailing the surface distribution of dissociation forces. (D) Statistical graph showing binding probabilities for different groups, highlighting significant differences determined by t-tests ($p<1e-5$). (E) Flow cytometry results illustrate interactions of wild-type (WT) and Uncharged H3 fused with eGFP with CHO-K1 cells, alongside surface HS removed control (green) and cell-only control (Blank, grey). (F) Statistical analysis of flow cytometry data, showing significant differences between groups as determined by t-tests. *****, $P<1e-5$

To explore the α-helical domain's role in binding HS, we engineered a variant of the H3 protein, named H3(uncharged), in which all positively charged residues were replaced with the uncharged amino acid serine. AFM experiments on this variant showed a detectable force peak with a lower unbinding force of 28.8 ± 0.2 pN, compared to 33.7 ± 0.3 pN for the wild type, along with a decreased HS binding probability from 43% to 25% (Fig. 4C, D). These findings indicate that the basic amino acids in the helical domain are crucial for HS binding. Additionally, to ensure the specificity of the interaction signals, we treated CHO-K1 cells with heparinase II, which degrades HS. Control experiments using this enzyme treatment demonstrated a further reduction in the dissociation force to 23.0 pN and a decrease in binding probability to 12.6%, confirming that the observed interactions specifically originated from the dissociation of the HS-H3 complex.

Further validation was carried out using flow cytometry (FCM). We fused eGFP, a fluorescent reporter protein, to the C-terminus of both the wild-type H3 (H3(WT)) and H3(uncharged). Following incubation of fusion proteins with cells, FCM analysis demonstrated a significant reduction in mean fluorescence intensity in cells treated with H3(uncharged) compared to H3(WT) (Fig. 4E). Statistical analysis of multiple experiments highlighted a significant difference between the two variants (Fig. 4F). These results reinforce the findings from the MD simulations and force spectroscopy, emphasizing that the positively charged amino acids within the helical domain of H3 are critical for the efficient binding to HS.

Design Protein Inhibitor for H3 Adhesion Targeting the Helical Domain

Following the discovery that the helical domain is crucial for the binding of H3 to HS, we focused on designing specific inhibitors to target this domain, aiming to both validate its functional significance and explore potential antiviral therapies using recently developed powerful AI tools for protein. Utilizing RFdiffusion, we designed

313 a series of inhibitors with α-helical structures and named it AI-PoxBlock (Fig. 5A).
314 Sequence recovery was performed on 1000 backbones using ProteinMPNN,
315 generating ten sequences per backbone. These designs were further validated through
316 complex structure predictions using AlphaFold2, selecting candidates based on pAE
317 values below 7 and iPTM scores above 0.9. Additionally, the predicted structures of
318 these candidates exhibited RMSD values less than 2 nm when compared to the
319 RFdiffusion-generated structures (Fig. S9). The interactions between the helical
320 domain of H3 and these inhibitors were then examined in 500 ns MD simulations,
321 confirming binding capabilities through the analysis of RMSD trajectories, where five
322 inhibitors—AI-PoxBlock302, 602, 614, 723, and 761—showed stable RMSD values
323 under 0.5 nm (Fig. S10, S11, Supplementary Video 6).
324 Although AI-PoxBlock302 could not be successfully expressed, the remaining
325 four inhibitors underwent further testing in FCM to evaluate their efficacy in
326 inhibiting H3 binding to CHO-K1 cell surfaces. AI-PoxBlock723 stood out,
327 significantly reducing H3 binding at a concentration of 10 μM, while the other
328 inhibitors displayed no tendency to inhibit H3 binding to HS on cell surfaces (Fig. 5B,
329 Fig. S12). Further analysis using Biolayer Interferometry (BLI) confirmed that AI-
330 PoxBlock723 binds to H3 with an equilibrium dissociation constant (K_D) of 9 μM
331 (Fig. 5C). To ensure specificity, we also assessed the interaction between AI-
332 PoxBlock723 and a truncated version of H3 (residues 1-239), lacking the α-helical
333 domain. BLI results indicated a marked reduction in binding, reinforcing the
334 specificity of AI-PoxBlock723 for the helical domain (Fig. S13). Circular Dichroism
335 (CD) spectroscopy confirmed that AI-PoxBlock723 predominantly consists of α-
336 helices, aligning with its design specifications (Fig. S14). To evaluate the antiviral
337 activity of AI-PoxBlocks, they were series diluted and incubated with MPXV before
338 the infection of Vero E6 cells and cultured overnight, followed by the determination
339 of virus infection foci as we previously described[46]. Excitingly, AI-PoxBlock723
340 showed a significant antiviral effect with the half maximal inhibitory concentration
341 (IC_{50}) of 8.86 μM, whereas other AI-PoxBlocks exhibited no inhibitory efficiency.
342 (Fig. 5D). Considering the high conservation of H3, particularly its helical domain
343 among the four *orthopoxviruses* pathogenic to humans (Fig. S15), we also assessed
344 the activity of AI-PoxBlocks against VACV infection of target cells. Consistently, AI-
345 PoxBlock723 is also effective for inhibiting VACA infection in vitro with a similar
346 IC50 of 8.67μM (Fig. S16), suggesting the helical domain as an antiviral drug pocket
347 is universal at least for *orthopoxvirus*. Collectively, these results demonstrate the
348 efficacy of the designed inhibitors targeting the H3 helical domain and validate the
349 crucial role of this domain in viral adhesion, providing a promising foundation for the
350 development of novel antiviral agents.
351

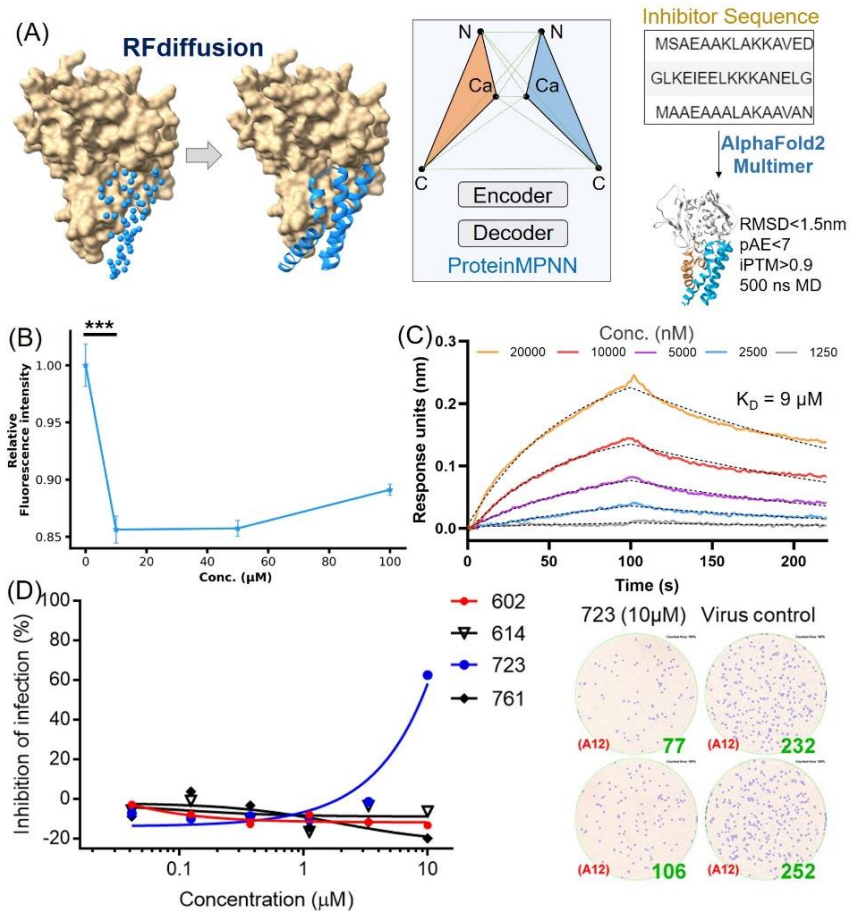


Fig. 5 **Development and Testing of Protein Inhibitors Targeting the H3 Helical Domain** (A) Diagrams depict protein inhibitors designed to target the H3 helical region, created using RFdiffusion. Sequences capable of folding into the target scaffold structures were generated using ProteinMPNN, and were validated through AlphaFold2, followed by 500 ns MD simulations for structural stability and interaction scoring. (B) FCM analysis demonstrates the inhibitory effect of AI-Poxblock723 at various concentrations. (C) BLI confirms direct interaction between AI-Poxblock723 and the H3 helical domain. (D) Graphs display the inhibitory effect of the indicated AI-Poxblocks on MPXV infection of Vero E6 cells, with quantitative analysis of virus-infected foci provided on the right.

Discussion:

Our study provides detailed mechanistic insights into the adhesion process of the MPXV, emphasizing the role of the H3 protein as a crucial adhesive component initiating host-cell interactions in the virus's lifecycle. By integrating computational methods with experimental validations, we have identified a previously unknown α-helical domain within the H3 protein that is essential for binding to HS on the host cell surface. This discovery enhances our understanding of MPXV pathogenesis and presents a novel target for therapeutic intervention.

HS serves as a cellular ligand for various viruses[47-49], underscoring the

importance of studying its interactions with viral adhesion proteins. However, the complexity and flexibility of polysaccharides like HS have historically posed challenges for the structural analysis of protein-polysaccharide complexes[48,50,51]. This complexity has rendered the molecular details of the interaction between MPXV's H3 and host cell HS elusive. By employing MD simulations and AI-based structural prediction tools such as AlphaFold2, we have overcome these obstacles, revealing key interaction sites—specifically the α-helical domain—and binding mechanisms[16,52-59]. Our study demonstrates the feasibility of our comprehensive computational approach in researching viral proteins involved in protein-glycan interactions, offering a promising methodological advancement.

Looking forward, the methodologies and findings from this study are poised to catalyze further investigations into *orthopoxviruses* and other viral families employing similar adhesion strategies. The development of innovative protein inhibitors, driven by state-of-the-art AI models for protein design including RFdiffusion, ProteinMPNN, and AlphaFold2, highlights the transformative impact of AI in drug discovery. Future research could focus on refining the efficacy and specificity of inhibitors targeting the α-helical domain and explore their clinical applicability[60]. Extending our integrative approach to other viral systems could reveal additional therapeutic targets, significantly enhancing our capacity to combat viral diseases.

In summary, our research provides a comprehensive computational methodology to study complex protein-glycan interactions and paves the way for the development of innovative antiviral drugs. These findings offer promising potential treatments for MPXV and other severe viral diseases, demonstrating the powerful capabilities of AI-assisted computational biology in the discovery of specific antiviral agents.

Acknowledgments: We acknowledge the funding support from the National Natural Science Foundation of China (22222703), the Natural Science Foundation of Jiangsu Province (BK20202004), Shenzhen Medical Research Funds (B2302052). The numerical calculations in this work have been done on the computing facilities in the High-Performance Computing Center (HPCC) of Nanjing University.

Author contributions:
Conceptualization: PZ
Investigation: BZ, MD, YX, SW, RG, LC
Funding acquisition: PZ, LC
Project administration & Supervision: PZ
Writing – original draft: PZ, BZ, LC

Competing interests: The authors declare no competing interests.

Data and materials availability: All data are available in the main text or the supplementary materials

Supplementary Materials: includes supplementary materials and methods, notes, figures and videos.

References:

1. Lum, F.-M. et al. Monkeypox: disease epidemiology, host immunity and

clinical interventions. *Nat. Rev. Immunol* **22**, 597-613 (2022).
2. The Lancet Infectious, D. Monkeypox: a neglected old foe. *Lancet Infect. Dis.* **22**, 913 (2022).
3. O'Toole, Á. et al. APOBEC3 deaminase editing in mpox virus as evidence for sustained human transmission since at least 2016. *Science* **382**, 595-600 (2023).
4. Kirby, T. What happened to the mpox pandemic? *The Lancet* **402**, 949-950 (2023).
5. Li, P. et al. Mpox virus infection and drug treatment modelled in human skin organoids. *Nat. Microbiol.* **8**, 2067-2079 (2023).
6. Liu, J. et al. Retrospective detection of monkeypox virus in the testes of nonhuman primate survivors. *Nat. Microbiol.* **7**, 1980-1986 (2022).
7. Zahmatyar, M. et al. Human monkeypox: history, presentations, transmission, epidemiology, diagnosis, treatment, and prevention. *Front. Med.* **10**(2023).
8. Zaeck, L.M. et al. Low levels of monkeypox virus-neutralizing antibodies after MVA-BN vaccination in healthy individuals. *Nat. Med.* **29**, 270-278 (2023).
9. Yu, J. et al. Phylogeny and molecular evolution of the first local monkeypox virus cluster in Guangdong Province, China. *Nat. Commun.* **14**, 8241 (2023).
10. Peng, Q. et al. Structure of monkeypox virus DNA polymerase holoenzyme. *Science* **379**, 100-105 (2023).
11. Moss, B. Investigating Viruses During the Transformation of Molecular Biology: Part II. *Annu. Rev. Virol.* **7**, 15-36 (2020).
12. Gray, R.D.M. et al. Nanoscale polarization of the entry fusion complex of vaccinia virus drives efficient fusion. *Nat. Microbiol.* **4**, 1636-1644 (2019).
13. Fonseca, F.G.d., Wolffe, E.J., Weisberg, A. & Moss, B. Effects of Deletion or Stringent Repression of the H3L Envelope Gene on Vaccinia Virus Replication. *J. Virol.* **74**, 7518-7528 (2000).
14. Crickard, L. et al. Protection of Rabbits and Immunodeficient Mice against Lethal Poxvirus Infections by Human Monoclonal Antibodies. *PLOS ONE* **7**, e48706 (2012).
15. Zuiani, A. et al. A multivalent mRNA monkeypox virus vaccine (BNT166) protects mice and macaques from orthopoxvirus disease. *Cell* **187**, 1363-1373.e12 (2024).
16. Gulati, P., Chadha, J., Harjai, K. & Singh, S. Targeting envelope proteins of poxviruses to repurpose phytochemicals against monkeypox: An in silico investigation. *Front. Microbiol.* **13**(2023).
17. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
18. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-802 (2005).
19. Gräter, F., Shen, J., Jiang, H., Gautel, M. & Grubmüller, H. Mechanically Induced Titin Kinase Activation Studied by Force-Probe Molecular Dynamics Simulations. *Biophy. J.* **88**, 790-804 (2005).

20. Kim, E. et al. A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nat. Chem. Biol.* **9**, 313-318 (2013).
21. Brooks, C.L., MacKerell, A.D., Post, C.B. & Nilsson, L. Biomolecular dynamics in the 21st century. *Biochim. Biophys. Acta. Gen. Subj.* **1868**, 130534 (2024).
22. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. & Gaub, H.E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109-12 (1997).
23. Sieben, C. et al. Influenza virus binds its host cell using multiple dynamic interactions. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 13626-13631 (2012).
24. Raab, A. et al. Antibody recognition imaging by force microscopy. *Nat. Biotech.* **17**, 901 (1999).
25. Müller, D.J., Helenius, J., Alsteens, D. & Dufrêne, Y.F. Force probing surfaces of living cells to molecular resolution. *Nat. Chem. Biol.* **5**, 383-390 (2009).
26. Dauparas, J. et al. Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **378**, 49-56 (2022).
27. Watson, J.L. et al. De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089-1100 (2023).
28. Singh, K. et al. The Vaccinia Virus H3 Envelope Protein, a Major Target of Neutralizing Antibodies, Exhibits a Glycosyltransferase Fold and Binds UDP-Glucose. *J. Virol.* **90**, 5020-5030 (2016).
29. Fonseca, F.G.d., Wolffe, E.J., Weisberg, A. & Moss, B. Characterization of the Vaccinia Virus H3L Envelope Protein: Topology and Posttranslational Membrane Insertion via the C-Terminal Hydrophobic Tail. *J. Virol.* **74**, 7508-7517 (2000).
30. Lin, C.-L., Chung, C.-S., Heine, H.G. & Chang, W. Vaccinia Virus Envelope H3L Protein Binds to Cell Surface Heparan Sulfate and Is Important for Intracellular Mature Virion Morphogenesis and Virus Infection In Vitro and In Vivo. *J. Virol.* **74**, 3353-3365 (2000).
31. Morgan, A. & Wang, X. The Novel Heparin-Binding Motif in Decorin-Binding Protein A from Strain B31 of Borrelia burgdorferi Explains the Higher Binding Affinity. *Biochemistry* **52**, 8237-8245 (2013).
32. Eberhardt, J., Santos-Martins, D., Tillack, A.F. & Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **61**, 3891-3898 (2021).
33. Lee, J. et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405-413 (2015).
34. Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141-151 (1999).
35. Govind Kumar, V., Polasa, A., Agrawal, S., Kumar, T.K.S. & Moradi, M. Binding affinity estimation from restrained umbrella sampling simulations. *Nat. Comput. Sci.* **3**, 59-70 (2022).

36. Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402 (1997).
37. Yu, H., Siewny, M.G.W., Edwards, D.T., Sanders, A.W. & Perkins, T.T. Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. *Science* **355**, 945-950 (2017).
38. Dietz, H. & Rief, M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16192-16197 (2004).
39. Goktas, M. et al. Molecular mechanics of coiled coils loaded in the shear geometry. *Chem. Sci.* **9**, 4610-4621 (2018).
40. Cao, Y., Lam, C., Wang, M. & Li, H. Nonmechanical protein can have significant mechanical stability. *Angew. Chem. Int. Ed.* **45**, 642-5 (2006).
41. Stahl, S.W. et al. Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20431-20436 (2012).
42. Deng, Y. et al. Enzymatic biosynthesis and immobilization of polyprotein verified at the single-molecule level. *Nat. Commun.* **10**, 2775-2785 (2019).
43. Carrion-Vazquez, M. et al. Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3694-3699 (1999).
44. Tian, F. et al. N501Y mutation of spike protein in SARS-CoV-2 strengthens its binding to receptor ACE2. *Elife* **10**, e69091 (2021).
45. Alsteens, D. et al. Nanomechanical mapping of first binding steps of a virus to animal cells. *Nat. Nanotechnol.* **12**, 177 (2016).
46. Cheng, L. et al. Isolation and characterization of mpox virus from the first mpox case in Shenzhen, China. *Virologica Sinica* **39**, 335-337 (2024).
47. Clausen, T.M. et al. SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell* **183**, 1043-1057.e15 (2020).
48. Shriver, Z., Capila, I., Venkataraman, G. & Sasisekharan, R. Heparin and Heparan Sulfate: Analyzing Structure and Microheterogeneity. in *Heparin - A Century of Progress* (eds. Lever, R., Mulloy, B. & Page, C.P.) 159-176 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012).
49. Koehler, M., Delguste, M., Sieben, C., Gillet, L. & Alsteens, D. Initial Step of Virus Entry: Virion Binding to Cell-Surface Glycans. *Annu. Rev. Virol.* **7**, 143-165 (2020).
50. Shih, P.-C. et al. A Turn-like Structure "KKPE" Segment Mediates the Specific Binding of Viral Protein A27 to Heparin and Heparan Sulfate on Cell Surfaces. *J. Bio. Chem.* **284**, 36535-36546 (2009).
51. Marszalek, P.E., Oberhauser, A.F., Li, H. & Fernandez, J.M. The Force-Driven Conformations of Heparin Studied with Single Molecule Force Microscopy. *Biophy. J.* **85**, 2696-2704 (2003).
52. Zhang, Z.Q., Xie, J., Liu, H.Y., Liu, J. & Linhardt, R.J. Quantification of Heparan Sulfate Disaccharides Using Ion-Pairing Reversed-Phase Microflow High-Performance Liquid Chromatography with Electrospray Ionization Trap Mass Spectrometry. *Anal. Chem.* **81**, 4349-4355 (2009).
53. Podduturi, S., Vemula, D., Singothu, S. & Bhandari, V. In-silico investigation

of E8 surface protein of the monkeypox virus to identify potential therapeutic agents. *J. Biomol. Struct. Dyn.*, 1-14 (2023).

54. Bernardi, R.C. et al. Mechanisms of Nanonewton Mechanostability in a Protein Complex Revealed by Molecular Dynamics Simulations and Single-Molecule Force Spectroscopy. *J. Am. Chem. Soc.* **141**, 14752-14763 (2019).
55. Zhao, M. & Woodside, M.T. Mechanical strength of RNA knot in Zika virus protects against cellular defenses. *Nat. Chem. Biol.* **17**, 975-981 (2021).
56. Serdiuk, T. et al. YidC assists the stepwise and stochastic folding of membrane proteins. *Nat. Chem. Biol.* **12**, 911-917 (2016).
57. Sen Gupta, P.S., Panda, S.K., Nayak, A.K. & Rana, M.K. Identification and Investigation of a Cryptic Binding Pocket of the P37 Envelope Protein of Monkeypox Virus by Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **14**, 3230-3235 (2023).
58. Petrosyan, R., Patra, S., Rezajooei, N., Garen, C.R. & Woodside, M.T. Unfolded and intermediate states of PrP play a key role in the mechanism of action of an antiprion chaperone. *Proc. Natl. Acad. Sci. U. S. A.* **118,** e2021213118 (2021).
59. Chen, G. et al. The adenylate cyclase toxin RTX domain follows a series templated folding mechanism with implications for toxin activity. *J. Biol. Chem.* **299**, 105150-105153 (2023).
60. Wang, J. et al. An Overview of Antivirals against Monkeypox Virus and Other Orthopoxviruses. *J. Med. Chem.* **66**, 4468-4490 (2023).

SUPPLEMENTARY MATERIALS

Materials and Methods

Protein expression and purification.

All genes were ordered from General Biosystems. H3(1-282) was constructed into the pCold-TF-tev vector using *BamH*I-*Bgl*II-*Kpn*I three-restriction enzyme system and a Strep-Tag II was added to the C-terminus for protein purification. eGFP was ligated to the C-terminus of H3(1-282) by restriction enzyme system. Mutations in the basic amino acids of the helical domain were introduced by site-directed PCR mutagenesis. The pCold-TF-tev-POI-strep tag construct was transformed into BL21(DE3), and single colonies were picked and cultured overnight at 37°C in LB medium with ampicillin. The expanded culture was then transferred to 800 mL of ampicillin LB and continued to be incubated at 37°C with shaking for 2 hours. When the OD_{600} of the culture reached 0.6-0.8, it was cooled to 15°C in an ice-water bath. Then, IPTG was added to a final concentration of 0.5 mM and incubated overnight at 15°C on a shaker.

The next day, the cells were collected by centrifugation at 8000 rpm, the supernatant was discarded, and the cell pellet was resuspended in buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4). The cell suspension was then lysed with a homogenizer for 3 minutes, and the lysate was centrifuged at 18000 rpm to separate the precipitate. TEV protease was added to the supernatant, along with EDTA and DTT to a final concentration of 1mM each, and incubated at 30°C for 3 hours to remove the TF tag. The digested product was subjected to another high-speed centrifugation at 18000 rpm, and the supernatant was collected. The supernatant was then purified using STarm Sreptactin.

The proteins used for AFM force spectroscopy were sequentially constructed into the pQE80L vector using the same restriction enzyme system, resulting in pQE80L-coh-(GB1)2-H3(1-282)-strep tag II-NAL. The protein expression method was similar to the one mentioned above. The difference was that after adding IPTG to a final concentration of 0.5 mM, the culture was incubated overnight at 18°C. After cell lysis and centrifugal removal of the precipitate, the proteins were purified using STarm Sreptactin. The eluate was exchanged into AFM working buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4) using an ultrafiltration tube and then used for AFM force spectroscopy experiments.

*Oa*AEP1(C247A) is cysteine 247 to alanine mutant of asparaginyl endoproteases 1 from *oldenlandia affinis*, abbreviated as AEP here. ELP is the elastin-like polypeptides (2). Their expression and purification protocols can be found in references. The TEV protease was expressed in the pCold-His6-ProS2 vector (the original tev protease cleavage site in the vector was removed by PCR). The expression was induced with a final concentration of 0.1mM IPTG and incubated

overnight at 15°C in a constant temperature shaker. The His6-ProS2-TEV was purified using Ni-NTA affinity chromatography.

All H3 inhibitors designed via RFdiffusion were constructed into the pET30a vector, with C-terminal fusion of strep tag and his8 tag for protein purification and subsequent BLI immobilization. After transforming pET30-Inhibitor-strep tag II-His8 into BL21(DE3), the culture was grown at 37°C in kanamycin-resistant LB medium for 2 hours until the OD_{600} reached 0.6-0.8, then IPTG was added to a final concentration of 0.5 mM and the culture was incubated overnight at 18°C. The subsequent purification steps were as previously described, using STarm Sreptactin. The purified protein was exchanged into PBS buffer via ultrafiltration.

Molecular Docking of H3 with HS

A HS oligosaccharide with the structure -[IdoA2S-GlcNS6S-IdoA-GlcNS(3,6S)]$_5$- was modeled using the CHARMM-GUI online website. The structure of H3(1-282) predicted by AlphaFold2 was refined through 500 ns of molecular dynamics simulation, and the most frequently appearing conformation was obtained through structural clustering. Specifically, by calculating the RMSD values of H3 in the MD trajectory, and then clustering using the built-in gmx cluster program, the gromos clustering method was adopted with a cut-off set to 0.2 nm. The top three conformations obtained were used for the subsequent docking process.

AutoDock Vina was called via the Chimera graphical interface to generate docking configuration files for HS and H3, with the center of each docking box set at the center of the GAGs binding motif, and the box size set to 8 nm. The energy_range was set to 3 kcal/mol, specifying that the energy difference from the current best conformation should be less than 3 kcal/mol. The docking results were ranked according to the energy scores generated by AutoDock Vina, and the best four conformations were selected.

MD Simulation of the H3-HS Complex

The docked H3-HS complex was uploaded to the CHARMM-GUI website[1] for simulation system construction, utilizing the charmm36m force field and the TIP3P water model. A water box, 1 nm larger than the molecular boundaries of the complex, was constructed, and 0.15 M NaCl was used for charge balancing. The final output included structure and force field files suitable for GROMACS 2023.3[2] MD simulations. After energy minimization to reduce the maximum force below 1000 kJ/mol/nm, an NVT equilibration was conducted using the V-rescale thermostat method at 310K for 1 ns. This was followed by a 1000 ns NPT production simulation at 310K using the V-rescale thermostat and the C-rescale barostat methods at 1 bar. Three independent repetitions were performed for the production simulation.

The RMSD values of HS were calculated using the built-in gmx rms program in GROMACS, maintaining the structural overlap of H3. To analyze salt bridges between HS and HS in this simulation, a Python script was used to calculate the distance between the negatively charged oxygen atoms on the side chains of HS and the nitrogen atoms on the side chains of basic amino acids of H3 throughout the

trajectory. A distance less than the set cutoff of 0.35 nm was considered indicative of salt bridge formation.

Binding Free Energy of H3 with HS Using Umbrella Sampling in MD Simulations

To calculate the stability of various binding conformations of H3-HS, the binding free energy was determined using the umbrella sampling method. The force field and solvent settings of the simulation system were consistent with the previously mentioned MD simulations, employing the CHARMM36m force field and TIP3 water model, and maintaining a simulation temperature of 310K. These simulations were conducted using GROMACS 2023.3. Specifically, for umbrella sampling, Steered Molecular Dynamics (SMD) simulations were first carried out. During this process, H3 was held fixed (using a restraining potential), and a harmonic potential was applied to the HS molecule. This potential facilitated the constant velocity stretching of HS (1 nm/ns) along the z-axis, moving it 5 nm until complete separation from H3 was achieved. The size of the simulation box was designed to ensure that neither molecule crossed the boundaries during stretching. The stretching process generated a series of reaction coordinates, defined by the distance between the centroids of HS and H3. Biasing potentials were applied to these coordinates for 10 ns MD simulations. The outputs of these simulations, in the form of potential energy distributions, were then analyzed using the Weighted Histogram Analysis Method (WHAM). This analysis produced a Potential of Mean Force (PMF) profile, describing the free energy landscape as a function of the reaction coordinate, thus allowing for the calculation of the binding free energy of H3-HS.

Replica Exchange Molecular Dynamics Simulation

To further investigate the conformational changes of HS and the helical domain, we conducted replica exchange MD simulations. The simulation system was set up using CHARMM-GUI, employing the same force field and water model as the aforementioned MD simulations, and 0.15 M NaCl was used for charge balancing. The temperature range was set from 310K to 387K, generating a total of 64 replicas with a swap probability of 20%. Each simulation had a duration of 1 μs, resulting in a cumulative simulation time of 64 μs for all replicas.

Simulations were executed using the MPI version of GROMACS 2023.3. Upon completion of the simulations, the replica at 310K was selected for further analysis. Using GROMACS built-in tools, 'distance' and 'rms', we calculated the nearest distance between HS and Mg(II), and the RMSD of HS throughout the simulation trajectory, respectively. Subsequently, the gmx sham program was used to generate a free energy landscape from these two sets of data. This landscape effectively illustrates the dynamic binding changes between HS and H3, as well as the interaction between HS and Mg(II).

***Poxviridae* Virus H3 Sequence Analysis**

A total of 66 available H3 sequences from *Poxviridae* viruses were obtained

through a sequence search of the monkeypox virus H3 sequence using NCBI's BLAST[3]. After removing duplicate sequence data, multiple sequence alignment was performed using MEGA11[4], with the ClustalW method[5] employed for sequence alignment. The alignment results were uploaded to the WebLogo website[6] for logo image generation. Acidic amino acids, basic amino acids, polar amino acids, and non-polar amino acids were represented in red, blue, green, and black, respectively. The font size in the logo corresponds to the probability of occurrence of the amino acid at that position, with larger fonts indicating greater conservation of the amino acid. The charge analysis of the amino acid sequences was conducted using the ProteinAnalysis function of the Biopython package. Surface charge maps of the H3 protein were generated using Chimera, and the charge distribution was calculated using APBS.

AFM-SMFS Unfolding Experiment.

The AFM cantilever/tip made of silicon nitride (MLCT-BIO-DC, Bruker Corp.) was used. The detailed protocol for AFM tip functionalization and protein immobilization on the glass coverslip can be found in references [7, 8]. In short, the tip and glass coverslip were coated with the amino group by amino-silanization. N_3 is functionalized on the surface from the reaction between $ImSO_2N_3 \cdot HCl$ and $–NH_2$. Then, a heterobifunctional DBCO-PEGn-Mal can be reacted and adds the Mal group. Next, the peptide GL-ELP_{20}-C or C-ELP_{20}-NGL was reacted to the maleimide via the cysteine, respectively. The long ELP_{20} serves as a spacer to avoid non-specific interaction between the tip and the surface as well as a signature for the single-molecule event. Finally, target protein H3 with C-terminal NGL sequence or GB1-Doc with N-terminal GL sequence can be site-specifically linked to the coverslip or tip by ligase AEP, respectively.

Atomic force microscope (Nanowizard4, JPK) was used to acquire the force-extension curve. The D tip of the MLCT-Bio-DC cantilever was used. Its accurate spring constant was determined by a thermally-induced fluctuation method [9]. Typically, the tip contacted the protein-immobilized surface for 100 *ms* under an indentation force of 350 pN to ensure a site-specifically interaction. Then, moving the tip up vertically at a constant velocity (1 μm/s), the polyprotein unfolded. Then, the tip moved to another place to repeat this cycle several thousands of times. As a result, a force-extension curve was obtained, which was analyzed using JPK data process analysis software.

AFM-SMFS Unbinding Experiment of H3-HS on CHO-K1

FD-based AFM on model surfaces was performed in PBS buffer at room temperature using functionalized D tip of MLCT-Bio-DC cantilever (Bruker, nominal spring constant of 0.030 N/m and actual spring constants calculated using thermal tune). AFM (Nanowizard4, JPK) operated in the force mapping (contact) mode was used. CHO-K1 cells were cultured in a 37°C CO_2 incubator for 24 hours prior to force spectroscopy experiments. The relative positioning of the AFM probe and the adherent cells was determined using an inverted microscope. Areas of 5 × 5 μm were scanned, ramp size set to 350 nm, and set point force of 300 pN with a contact time of

200 ms, with a resolution of 32 × 32 pixels.

Flow Cytometry Experiment

Protein binding to glycosaminoglycans (GAGs) expressed on the surface of Chinese Hamster Ovary K1 (CHO-K1) cells was assessed using flow cytometry, following the previously described methods. H3-eGFP, H3(uncharged)-eGFP, and eGFP proteins (5 μM) were incubated with 1 million cells (200 μL) at 4°C for 20 minutes. Post-incubation, cell analysis was performed using the CytoFLEX flow cytometer (Beckman) to collect fluorescence signals in the FITC channel. A total of 15,000 events were collected using the CytExpert software (version 2.5; Beckman), with 10 sets of data collected in parallel. The data from the live-cell gating were analyzed using FlowJo software (version 10.6.2; Tree Star, San Carlos, CA).

The flow cytometry experiment to validate the inhibitory effects of inhibitors utilized the same equipment and methods as previously described. The final concentrations of H3-eGFP and the inhibitor were 2 μM and 10 μM, respectively, with a total volume of 600 μL. The control group contained only 2 μM H3-eGFP, with the volume made up with PBS. Data from five parallel experiments were collected, with each experiment gathering 15,000 events. Using FlowJo, the fluorescence intensity of the mean of FITC-area signal in the live cell gate was analyzed, normalizing the relative fluorescence intensity of the control group to 1.

Design of Inhibitors Targeting the H3 Helical Domain

To design inhibitors targeting the helical domain of H3, hotspots information was first inputted into RFdiffusion. The selected residues for this purpose were 239, 242, 267, 266, 259 of H3, and the length of the inhibitor sequences was set to be 40-80 amino acids. This process generated 1000 backbone structures. After excluding single-stranded helical structures, the remaining 633 structures underwent sequence recovery using ProteinMPNN, producing 10 sequences per structure. All 6330 sequences were then subjected to structure prediction and scoring using AlphaFold2. Structures with an iPTM score greater than 0.9 and a pAE score less than 6 were selected, while sequences containing Cys were excluded to avoid the formation of disulfide bonds.

RMSD calculations were performed between the predicted structures and the RFdiffusion-designed structures, and sequences with an RMSD greater than 1.5 nm were excluded to ensure consistency between the predicted and designed structures. Additionally, solubility assessments were conducted for all sequences. Specifically, a solubility parameter was assigned to each of the 20 natural amino acids, and an average was calculated for the entire sequence. The sequences with the highest solubility were selected for expression and interaction testing.

MD Simulation of H3 and its Inhibitors.

he AlphaFold2-predicted structure of the H3-inhibitor complex was used for the MD simulation, utilizing the previously mentioned MD software and force field. The MD simulation was conducted at 310K for 500 ns. Subsequently, the RMSD of the

simulation trajectory was calculated using the built-in 'rms' program of GROMACS to analyze the stability of the inhibitor-H3 binding.

Circular Dichroism Experiments

For circular dichroism (CD) experiments, AI-Poxblock723 were diluted to 0.15mg/ml in 10 mM K-PO_4 (pH 7.4) buffer. Spectra were acquired on a Chirascan V100 (Applied Photophysics). The data acquisition wavelength is set to 190-260 nm. All reported measurements were acquired within the linear range of the instrument.

Bio-layer Interferometry (BLI) Binding Experiments

BLI experiments were conducted using Octet HIS1K Biosensors on the Octet system. The buffer used throughout the experiments was uniform PBS. The tips were first pre-equilibrated in PBS solution for 10 minutes, followed by a 60 s baseline, 60 s loading, 100 s baseline, 100 s association, and 120 s dissociation steps. Baseline measurements of unloaded tips were subtracted from their matched measurement of the loaded tip. The inhibitor concentration was determined using the BCA method before loading and prepared at a concentration of 0.1 mg/mL for the loading solution. The mass concentration of H3(1-282) and H3(1-239) was also determined using the BCA method, and their molar concentrations were calculated. Various concentration samples were then prepared using a serial dilution method. After baseline correction and curve smoothing, a global fit was performed to calculate the K_D.

Inhibition Test for Viral Infection

Vero E6 cells were seeded at 1.5×10^4 cells/well into 96-well plates and used the following day. AI-Poxblocks 3-fold serial diluted in maintenance medium were mixed with an equal volume of diluted live VACV or MPXV and then incubated at 37 °C for 1 hour. Medium from 96-well plates was aspirated, and the inhibitor-virus mixture was added (100 μl/well), then the Vero E6 cells were incubated at 37 °C for about 16 hours. Then cells were fixed with 4% paraformaldehyde solution, permeabilized with Perm/Wash buffer (BD Biosciences) containing 0.1% Triton X-100, incubated with the HRP-conjugated anti-VACV polyclonal antibodies (Invitrogen) diluted in the Perm/Wash buffer at room temperature for 2 hours. The reactions were developed with KPL TrueBlue Peroxidase substrates (Seracare Life Sciences). The numbers of infected foci were calculated using an EliSpot reader (Cellular Technology Ltd). The 50% inhibitory concentration (IC50) was calculated using GraphPad Prism software using the log (inhibitor) vs. normalized response-variable slope (four parameters) model. Experiments with live MPXV were performed in a Biosafety Level 3 (BSL-3) facility following standard biosafety practices.

Supplementary Text

Protein sequences:

H3 unfolding

His6-Coh-GB1-GB1-H3(1-282)-strep tag II-NGL
MRGSHHHHHHGSMGTALTDRGMTYDLDPKDGSSAATKPVLEVTKKVFDTAA
DAAGQTVTVEFKVSGAEGKYATTGYHIYWDERLEVVATKTGAYAKKGAALE
DSSLAKAENNGNGVFVASGADDDFGADGVMWTVELKVPADAKAGDVYPID
VAYQWDPSKGDLFTDNKDSAQGKLMQAYFFTQGIKSSSNPSTDEYLVKANAT
YADGYIAIKAGEPRSMDTYKLILNGKTLKGETTTEAVDAATAEKVFKQYAND
NGVDGEWTYDDATKTFTGTERSMDTYKLILNGKTLKGETTTEAVDAATAEK
VFKQYANDNGVDGEWTYDDATKTFTGTERSMAAVKTPVIVVPVIDRPPSETF
PNVHEHINDQKFDDVKDNEVMQEKRDVVIVNDDPDHYKDYVFIQWTGGNIR
DDDKYTHFFSGFCNTMCTEETKRNIARHLALWDSKFFTELENKNVEYVVIIEN
DNVIEDITFLRPVLKAIHDKKIDILQMREIITGNKVKTELVIDKDHAIFTYTGGY
DVSLSAYIIRVTTALNIVDEIIKSGGLSSGFYFEIARIENEMKINRQIMDNSAKYV
EHDPRLVAEHRFETMKPNFWSRIGTVAAKRYPGVMYTFTTPLISFRSWSHPQF
EKGSNGL

H3 unbinding on CHO-K1 cells

The following proteins were obtained by cleaving his6-TF-tev-POI using TEV protease.

H3(1-282)-strep tag II-NGL
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV
VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAA
KRYPGVMYTFTTPLISFRSWSHPQFEKGSNGL

H3(uncharged)-strep tag II-NGL
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV
VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPSLVAESSFETMSPNFWSSIGTVAAS
SYPGVMYTFTTPLISFRSWSHPQFEKGSNGL

Flow Cytometry Experiment

The following proteins were obtained by cleaving his6-TF-tev-POI using TEV protease.

H3(1-282)-eGFP- strep tag II
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV

VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAA
KRYPGVMYTFTTPLISFRSPAPAPAPMVSKGEELFTGVVPILVELDGDVNGHKF
SVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKED
GNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQ
NTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELY
K-RSWSHPQFEK

H3(uncharged)-eGFP- strep tag II
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV
VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPSLVAESSFETMSPNFWSSIGTVAAS
SYPGVMYTFTTPLISFRSPAPAPAPMVSKGEELFTGVVPILVELDGDVNGHKFS
VSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKED
GNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQ
NTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELY
K-RSWSHPQFEK

eGFP- strep tag II
GGSMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICT
TGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFK
DDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIM
ADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQS
ALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK-RSWSHPQFEK

Inhibitor design

AI-Poxblock602-strep tag II-His8:
MSAEAAKLAKKAVEDPSYAEELLKKDAESSTEERAAIANALLAERRRNPEKA
QKMIEKAAKIAFEDAKKELEEEEKKKAERSGGGGSGGGGSWSHPQFEKGGS
HHHHHHHH
MW: 12.14 kDa

AI-Poxblock614-strep tag II-His8:
MGLSKLNELKAKANALGKEARAALDAGDFEKAKEKILAGAKATKEAGELSG
NKAMIEDGEKAEEVAEEIVKIAKEERAKERSGGGGSGGGGSWSHPQFEKGGS
HHHHHHHH
MW: 11.68 kDa

AI-Poxblock723-strep tag II-His8:
MGLKKLQELKKKANELGKKAKEALDKGDFEKAKEYIKKGAEATKEYGELSG
NKAAIEDGEKAEEVAEEIVKIAKEEKAKERSGGGGSGGGGSWSHPQFEKGGS
HHHHHHHH
MW: 12.05 kDa

AI-Poxblock761-strep tag II-His8:
MSEFLESLKKAEELRKEVRELMEKGKEEADKLYKEGKEEEAAKVFLETAKEA
EPIAEEALKLLMERSGGGGSGGGGSWSHPQFEKGGSHHHHHHHH
MW: 10.74 kDa

The following proteins were obtained by cleaving TF-tev-POI-strep with TEV protease. To avoid the influence of the His tag from the pCold vector on BLI assays, site-directed mutagenesis via PCR was employed to remove the His tag present in the vector.

H3(1-282)-strep tag II-NGL
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV
VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAA
KRYPGVMYTFTTPLISFRSWSHPQFEKGSNGL

H3(uncharged)-strep tag II-NGL
GGSMAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEVMQEKRDV
VIVNDDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARH
LALWDSKFFTELENKNVEYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREI
ITGNKVKTELVIDKDHAIFTYTGGYDVSLSAYIIRVTTALNIVDEIIKSGGLSSGF
YFEIARIENEMKINRQIMDNSAKYVEHDPSLVAESSFETMSPNFWSSIGTVAAS
SYPGVMYTFTTPLISFRSWSHPQFEKGSNGL

Supplementary Figures

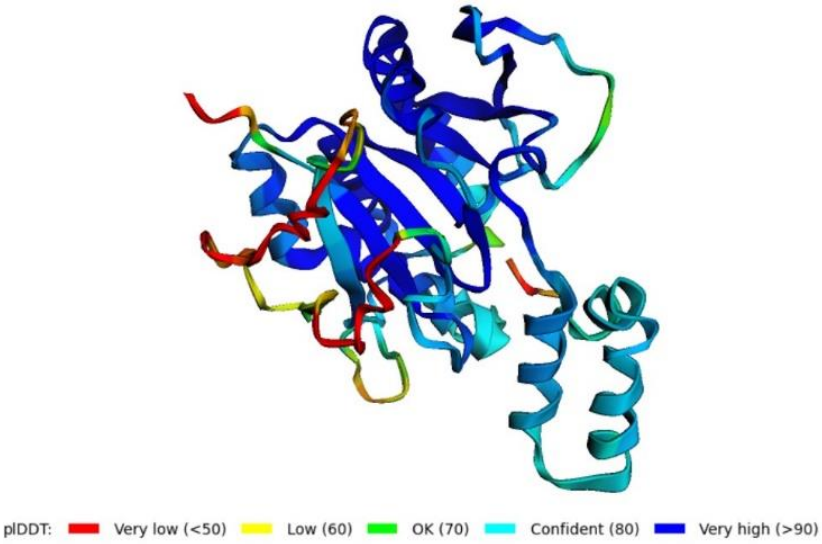
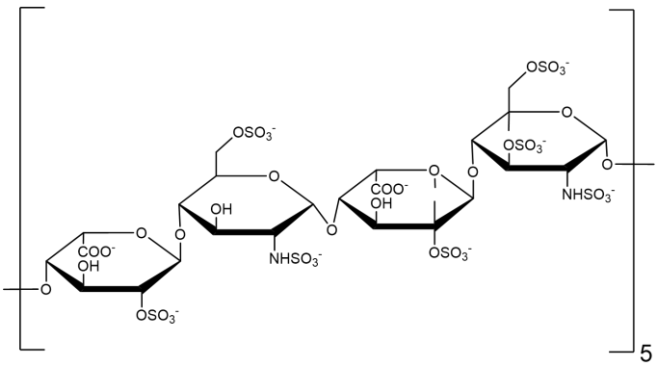


Fig. S1 Structural Prediction of Monkeypox Virus H3 Protein by AlphaFold2. In the protein cartoon diagram, the colors represent the pLDDT scores output by AlphaFold2, with blue indicating a higher prediction confidence. The helical domain is displayed in blue, indicating a region where the pLDDT score is greater than 80, signifying a high confidence in structural prediction.



-[IdoA2S-GlcNS6S-IdoA2S-GlcNS(3,6S)]$_5$-

Fig. S2 Structural Formula of HS. The composition of HS used in the main text consists of -[IdoA2S-GlcNS6S-IdoA-GlcNS(3,6S)]$_5$-, containing a total of 20 repeating monosaccharide units.

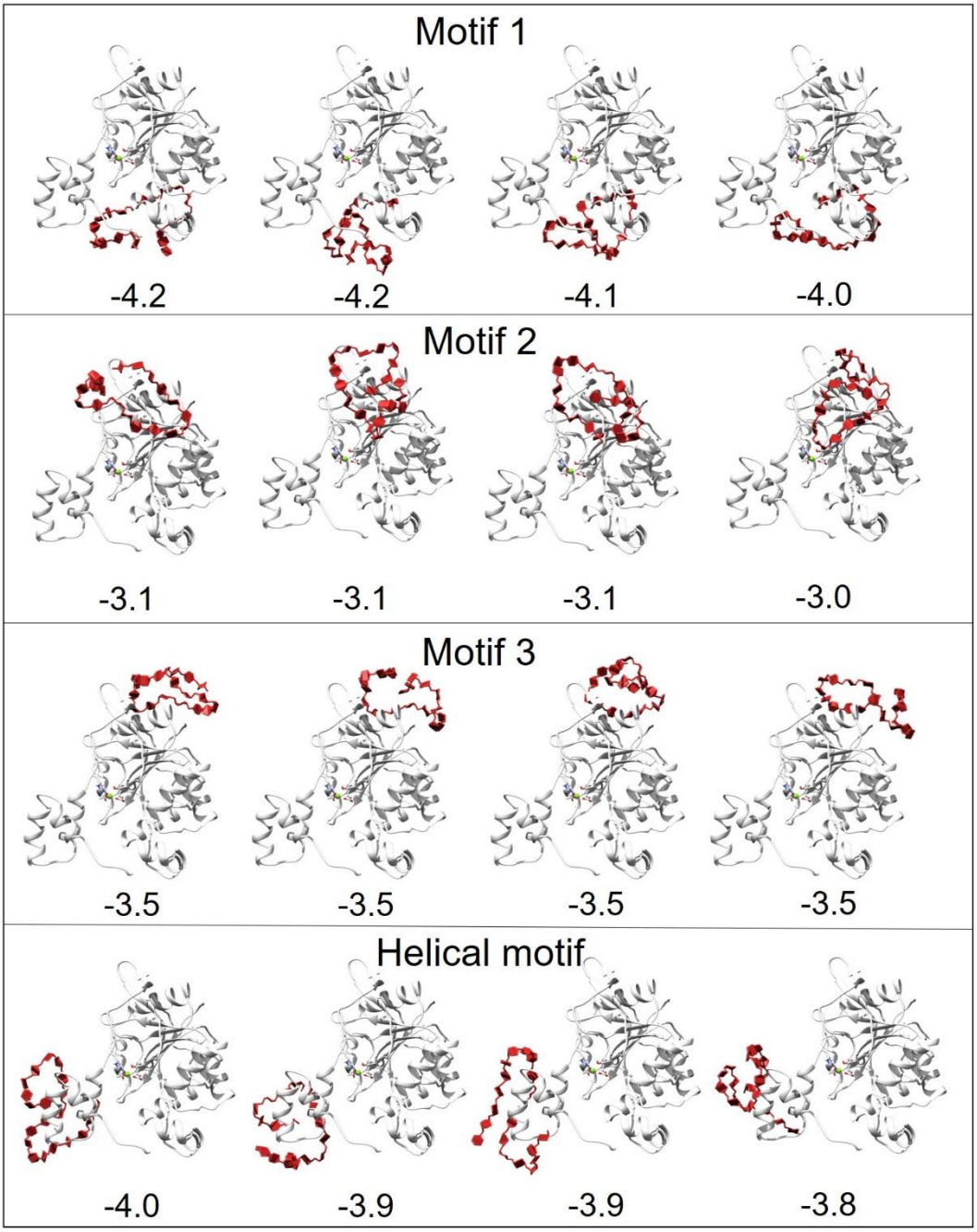


Fig. S3 Docking Results of H3 with HS. The docking area of HS was confined near different motifs by setting the docking box. The figure displays the docking results for four motifs, with the top four docking outcomes in each motif area selected based on the best scores (kCal/mol) from AutoDock Vina for subsequent MD simulations.

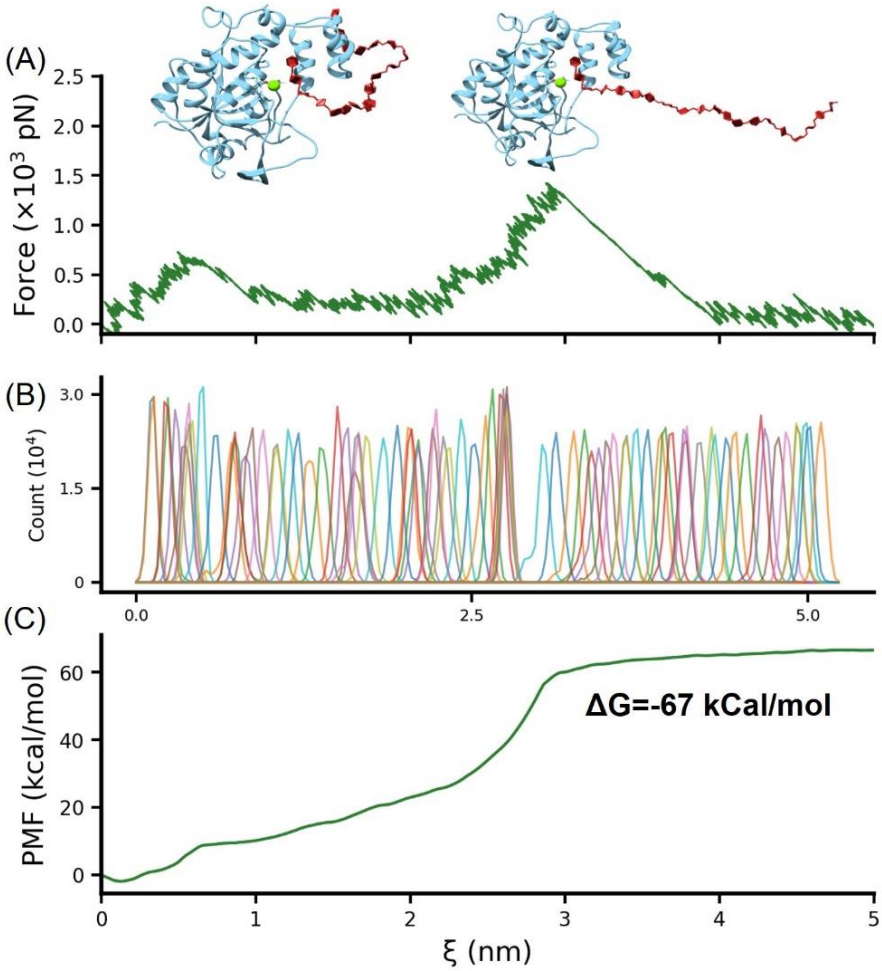


Fig. S4 Umbrella Sampling Calculation of the Binding Free Energy of HS-H3. The conformation of HS binding simultaneously to the helical domain and Mg^{2+}, obtained from replica exchange simulations, was further subjected to umbrella sampling to calculate the binding free energy. The stretching process is illustrated in (A), with the distribution of the reaction coordinate shown in (B). (C) presents the PMF (Potential of Mean Force) profile obtained after calculations using the Weighted Histogram Analysis Method (WHAM).

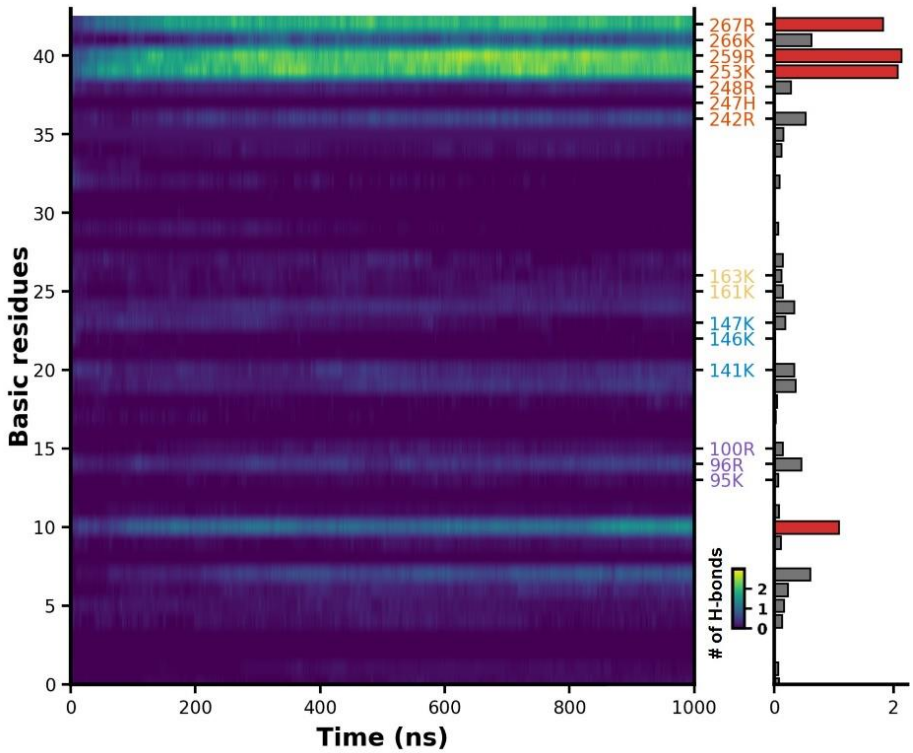


Fig. S5 Evolution of Salt Bridge Formation in HS During MD Simulations. This represents the variation in the number of salt bridges formed between HS and all basic amino acid side chain of H3 as the simulation progresses. The lighter areas indicate the formation of a greater number of salt bridges.

Fig. S6 AlphaFold2 Structural Prediction of 66 *Poxviridae* Virus H3 Proteins. The yellow portion represents the helical domain, while the blue portion signifies the main body of H3.

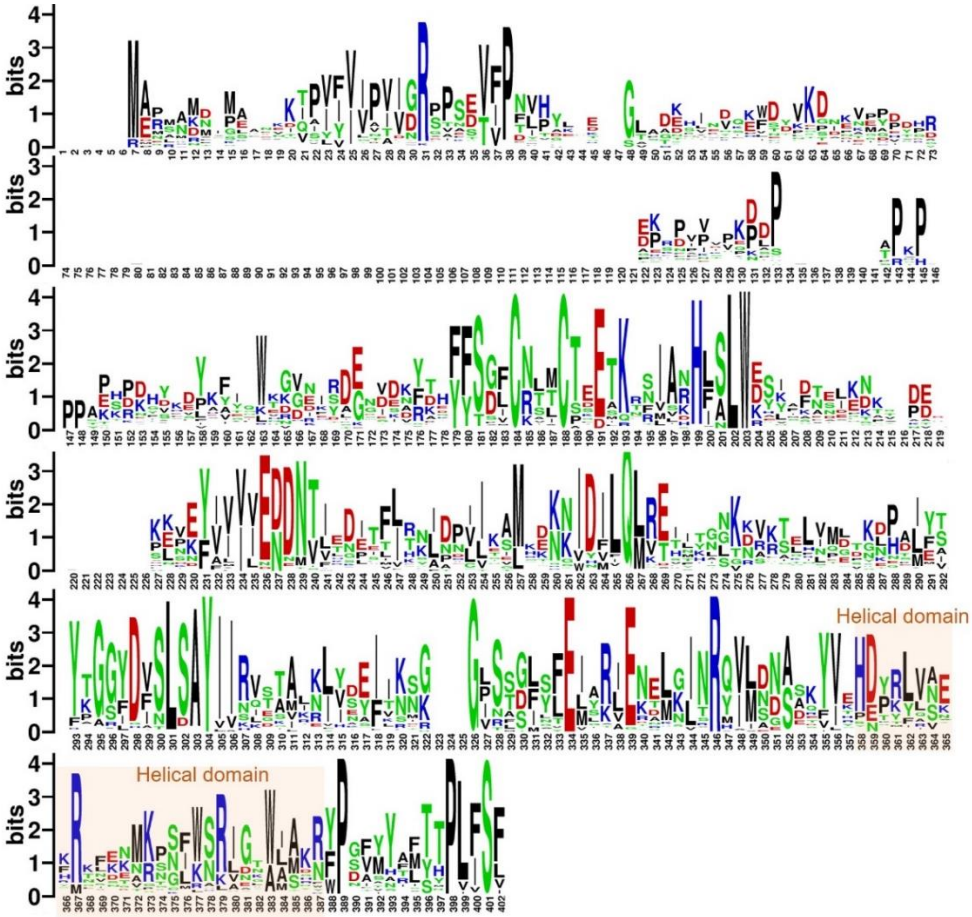


Fig. S7 Analysis Results of Full Sequence Analysis for 66 *Poxviridae* Virus H3 Proteins. The logo plot illustrates the probability of amino acid occurrence within the H3 sequences, where larger font sizes indicate a higher occurrence probability, denoting greater conservation. Acidic amino acids, basic amino acids, polar amino acids, and non-polar amino acids are represented in red, blue, green, and black, respectively.

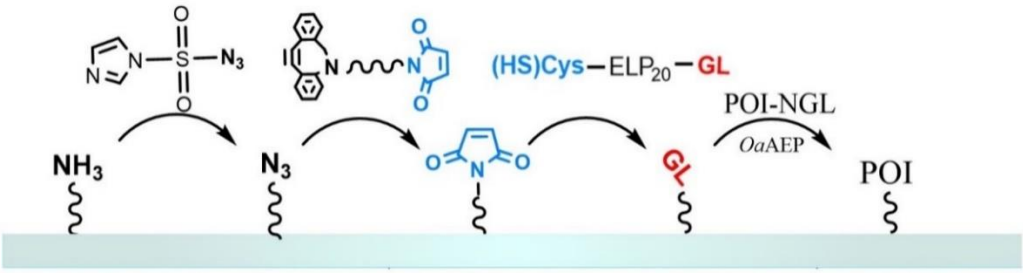


Fig. S8 Protein Immobilization in AFM-SMFS Experiments. Peptide GL-ELP_{20}-Cys is modified on a glass substrate (or AFM tip) that has been functionalized with maleimide through a Michael addition reaction. The N-terminus GL can undergo an enzymatic ligation reaction with the C-terminus NGL of the target protein in the presence of ligase *Oa*AEP1, thereby immobilizing the target protein on the substrate or AFM tip.

Fig. S9 Artificial Design of H3 Inhibitors. The gray structure represents the protein backbone structure designed by RFdiffusion, while the purple structure is the sequence generated by ProteinMPNN based on that backbone, followed by complex prediction results from AlphaFold2. The two structures align well, with an RMSD less than 1.5 nm.

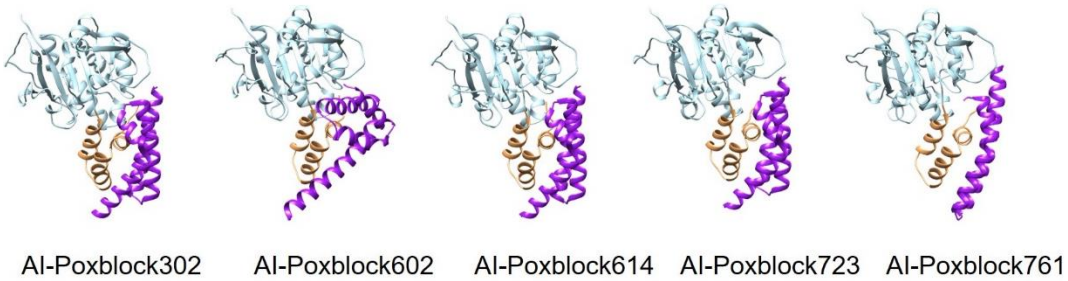


Fig. S10 AlphaFold2-predicted structures of the H3 and inhibitor complexes. From left to right are AI-Poxblock302, AI-Poxblock602, AI-Poxblock614, AI-Poxblock723, and AI-Poxblock761. All inhibitors are depicted in purple, and the helical domain of H3 are marked in yellow.

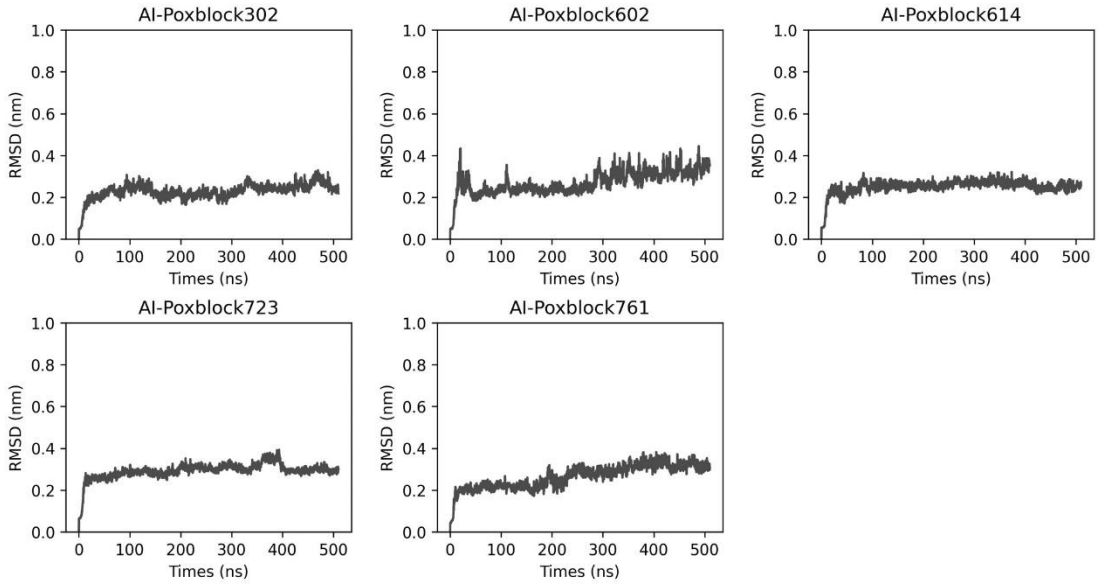


Fig. S11 MD simulations of H3-Inhibitor complex. The H3 and Inhibitor complex underwent a 500 ns MD simulation, resulting in the variation of RMSD (nm) over time (ns).

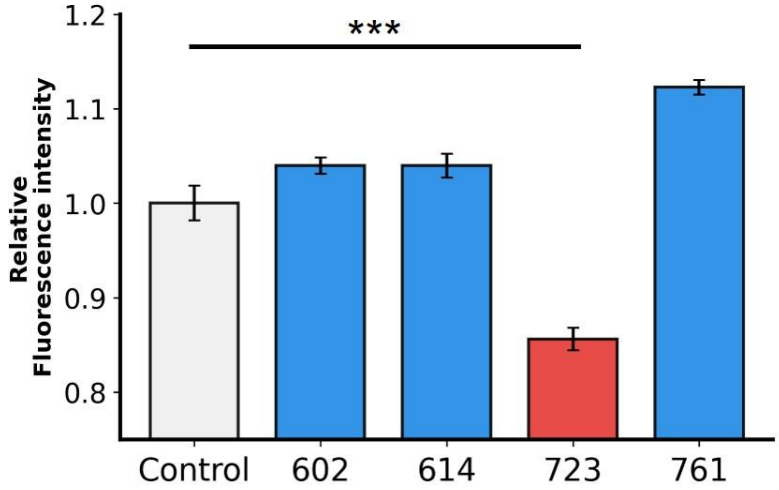


Fig. S12 Flow cytometry analysis demonstrates the efficacy of the inhibitors. The graph shows relative fluorescence intensity, indicating that in the presence of 10 μM AI-Poxblock723, the binding of H3(1-282)-eGFP to CHO-K1 cells is reduced compared to the control group (gray).

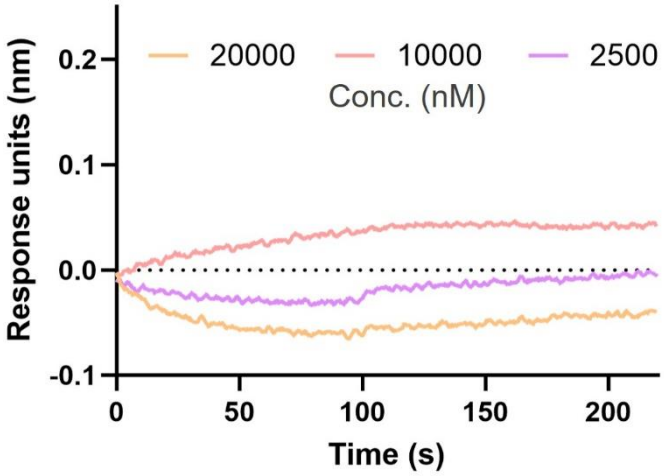


Fig. S13 AI-Poxblock723 and H3(1-239) BLI experiment. BLI curves of three concentrations of H3(1-239) (20000 nM, 10000 nM, and 2500 nM) with AI-Poxblock723 loaded on the sensor. The curves could not be fitted to obtain the k_D value.

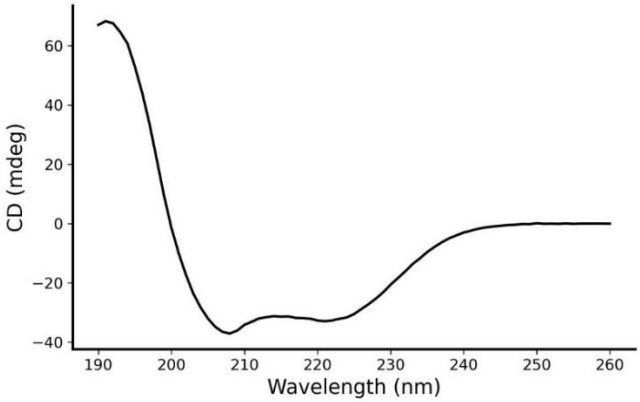


Fig. S14 CD spectroscopy experiment of AI-Poxblock723. CD spectroscopy verifies the designed inhibitor with a α-helical structure.

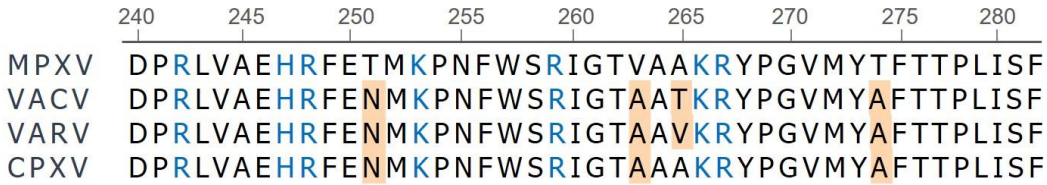


Fig. S15 Sequence comparison of the H3 helical domain among representative *orthopoxvirus* strains MPXV, VACV, VARV, and CPXV. Basic amino acid residues are highlighted in blue, and differences relative to the MPXV sequence are shaded in yellow.

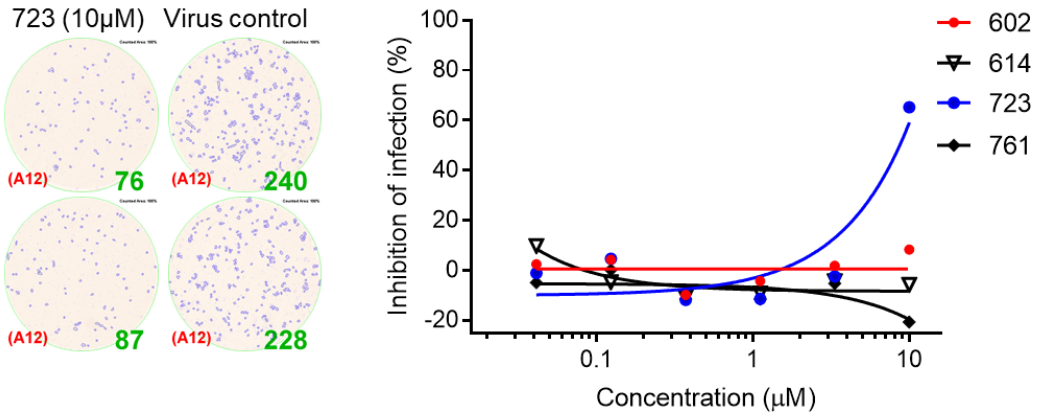


Fig. S16 The inhibitory effect of indicated inhibitors on VACV infection of Vero E6 cells. Virus-infected foci and counts were shown on the left.

Supplementary Video:

Supplementary Video 1: MD simulation trajectories of HS binding to H3 Domain 1 and the Helical domain.

Supplementary Video 2: SMD simulation trajectories of HS dissociating from H3 in umbrella sampling.

Supplementary Video 3: Trajectories of H3 and HS interactions in replica exchange MD simulations (REMD).

Supplementary Video 4: MD simulation trajectories of the H3-H3 complex after removal of the Mg^{2+} from H3.

Supplementary Video 5: MD simulation trajectories of the H3-H3 complex after mutating the positively charged amino acids on the helical domain of H3 to serine.

Supplementary Video 6: MD simulation trajectories of the AI-Poxblock723 complex.

Supplementary References:

1. Jo, S., Kim, T., Iyer, V.G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859-1865 (2008).
2. Abraham, M.J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25 (2015).
3. Altschul, S.F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).
4. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* **38**, 3022-3027 (2021).
5. Thompson, J.D., Higgins, D.G. & Gibson, T.J. Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research* **22**, 4673-4680 (1994).
6. Crooks, G.E., Hon, G., Chandonia, J.M. & Brenner, S.E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190 (2004).
7. Deng, Y. *et al.* Enzymatic biosynthesis and immobilization of polyprotein verified at the single-molecule level. *Nat. Commun.* **10**, 2775-2785 (2019).
8. Shi, S. *et al.* Combination of Click Chemistry and Enzymatic Ligation for Stable and Efficient Protein Immobilization for Single-Molecule Force Spectroscopy. *CCS Chem.* **4**, 598-604 (2022).
9. Hutter, J.L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Revi. Sci. Instrum.* **64**, 1868-1873 (1993).