Kirolos Makar

Title: Computational Evaluation of *Ebolavirus* VP35 in Viral Persistence: Evidence from Sequence and Structural Conservation

Abstract:

Ebolavirus is an extremely pathogenic virus and can persist in immune-privileged tissues for long periods of time even after an individual has recovered from illness; the biochemical basis for persistence of the virus is not yet clear. However, one of the key viral proteins, VP35, is known to play a role in replication and evasion of the immune response during acute infection; thus, it is reasonable to speculate that VP35 may also play an important role in the long-term persistence of the virus. The main objective of this study was to determine if VP35 retains structural and functional integrity during persistence, and whether such retention could support VP35 as a factor in maintaining *Ebolavirus* survival after recovery. To accomplish this, a variety of computational analyses were performed to compare the VP35 protein from acute versus persistent isolates. These analyses included BLAST, multiple sequence alignments, DALI structural analysis, PyMOL interface modeling, alanine scanning, and InterPro functional domain analysis. The results clearly demonstrate that VP35 is highly conserved at both the sequence level (100% sequence identity of VP35 between acute and persistent isolates) as well as at the structural level (nearly identical three-dimensional structures of VP35). The functional domain analysis also confirms that important regions of VP35 that are involved in the inhibition of the interferon response and the binding of RNA are preserved, while alanine scanning identified specific amino acids needed for VP35 to bind stably to the dynein light chain of the

host. Therefore, these data indicate that VP35 does not undergo mutations during persistence; rather, it maintains its biochemical sequence and structure during persistence. Therefore, it suggests that VP35 may play a role in the persistence of *Ebolavirus* by maintaining its ability to evade the immune response rather than through mutations. These findings contribute to our understanding of the mechanisms involved in viral persistence and highlight VP35 as a potential therapeutic target to prevent the long-term survival of the *Ebolavirus*. However, additional laboratory experiments will be necessary to validate these computational predictions.

Introduction:

Ebolavirus is an extremely pathogenic virus that causes severe hemorrhagic fever and results in death for many of those infected. Although the first phase of Ebola infection includes a rapid growth of the virus (viral replication) and severe inflammation, it has an additional feature to hide or persist in immune-isolated sites after recovery. Ebola can persist in the testes, ocular fluid, and cerebrospinal fluid and can hide in these areas for months to years after recovery.[1] Additionally, Ebola can emerge from a survivor and cause new outbreaks. Despite the importance of persistence, the biochemical mechanism behind it is unknown.[2]

One of Ebola's key proteins during the acute infection phase is Virus Protein 35 (VP35) which performs several functions. The first critical function of VP35 during the acute infection phase is to support the viral RNA dependent polymerase complex as a polymerase complex as a polymerase cofactor enabling replication of the viral genome. The second critical function of VP35 during infection is to antagonize the innate immune system. It specifically inhibits interferon production by binding the double strand RNA and preventing recognition receptors such as RIG-I and MDA5.[3] Without the interferon signaling, the virus does not get detected by the immune system allowing rapid viral replication. Also, VP35 interacts with host proteins such as dynein light chain (LC8) which is a protein involved in intracellular transport, and this interaction helps stabilize VP35 structure and may help in positioning viral components within the host cell which makes it important for viral replication and immune evasion.

The *Ebolavirus* protein VP35 is a critical factor in this regard because it prevents the activation of one of the first cell types to detect virus particles; this is the type I interferon response. When a virus infects a host, it normally produces viral RNA or double-stranded RNA replication intermediates, which would typically stimulate cytoplasmic pattern-recognition receptors (PRRs); specifically RIG-I and MDA5.[3] Once activated, they activate the mitochondrial antiviral signaling protein (MAVS) and other downstream kinases that activate transcription factors such as IRF-3 and IRF-7. Once activated, these transcription factors enter the nucleus and stimulate the production of type I interferons particularly IFN-α and IFN-β. The type I interferons warn other cells in the surrounding area and induce the expression of antiviral genes that inhibit the ability of the virus to replicate. Therefore, during the infection of the host with the *Ebolavirus*, if the VP35 protein is still conserved throughout the persistence of the virus, the viral particle would continue to be somewhat hidden from the host's immune system.

VP35 is a multi-functional 340-amino-acid protein with several different functional regions, each of which is associated with a separate viral function. The N-terminal region of VP35 is involved in oligomeric formation of VP35, binding with other viral or host factors. The C-terminal region of VP35 contains the interferon inhibitory domain (IID) which binds dsRNA and inhibits recognition by the innate immune system.[4] The IID is the critical site for investigation using computational methods since conservation of the IID would suggest that VP35 continues to have the necessary structural features that allow it to evade immune detection. Structural studies of the IID of VP35 have indicated that the IID has a unique folded domain and that a number of conserved basic residues throughout the IID are critical to binding dsRNA and inhibiting the effects of interferons. Additional structural studies of VP35 have indicated that it binds to dsRNA in such a way that it partially covers the ends and the backbone of dsRNA, which explains how VP35 inhibits the recognition of dsRNA by RIG-I and MDA5.[4] This makes the IID a key region to examine computationally because conservation of this domain would suggest that VP35 maintains the structure needed for immune evasion.

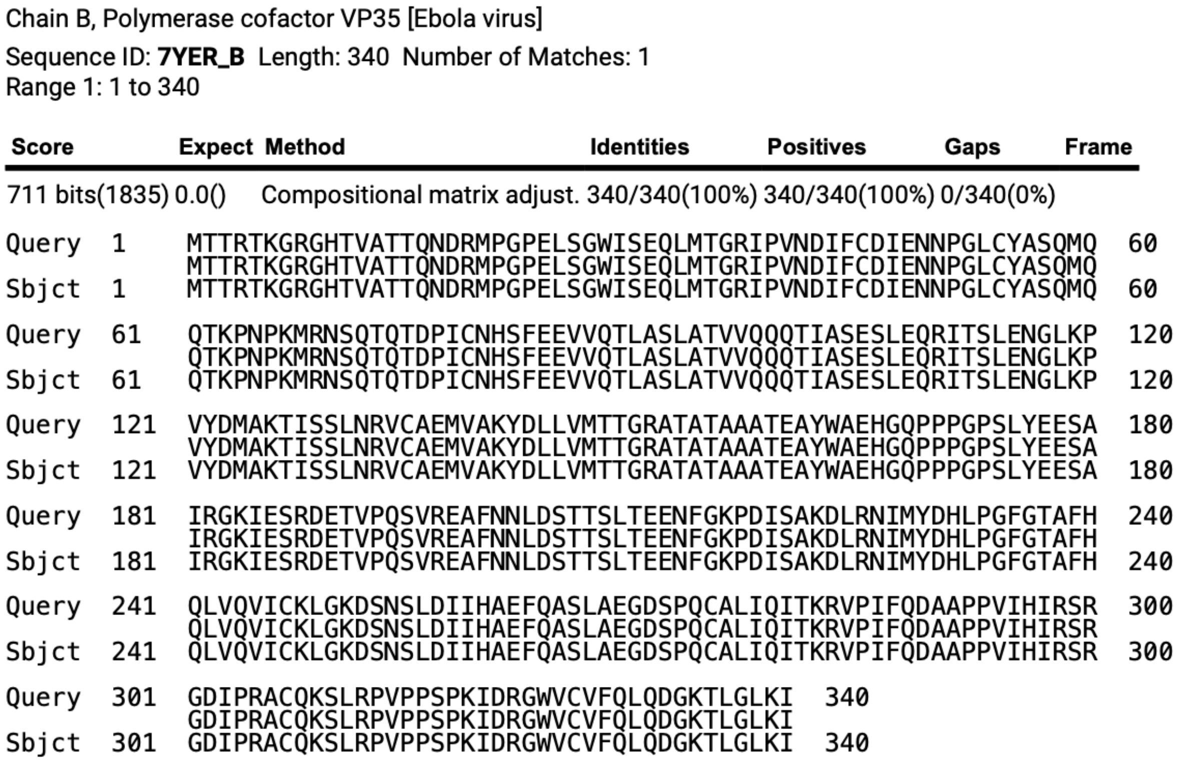
Importantly, the same immune evasion mechanisms during acute infection may also contribute to viral persistence after recovery. Since the persistence requires the virus to be undetected by the immune system, the proteins that play a role in evading the immune system contribute to persistence. The role of computational analysis in this study will be beneficial for analyzing VP35 at the sequence and structural level while rapid and efficient. Since viral persistence is long term and has an unknown mechanism, the computational analysis will enable us to predict which regions of VP35 are important and conserved. By comparing both sequence and structure of VP35 during infection and persistence, the conserved residues that may be important can be located. Therefore, this study uses computational analysis including BLAST and Multiple Sequence Alignment to evaluate sequence conservation, DALI structure alignment to assess structural similarity, PyMOL to visualize protein interactions, InterPro to identify functional domains, and alanine scanning to determine the energetic importance of key residues. These analyses were selected to analyze if VP35 maintains its structure and function during persistence which is essential to determine its potential role in viral persistence.

Sequence Conservation Analysis of VP35 using BLAST:

To evaluate if VP35 changes during viral persistence, the full length 340 amino acid VP35 sequence from a persistence isolate (ARG43180.1, from semen of Ebola survivor) was analyzed using BLASTp using the NCBI non reductant protein database.[5] There was complete alignment between persistence derived VP35 and reference *Ebolavirus* VP35 (7YER_B) across all the 340 amino acids from the BLAST results (Fig. 1), and the top structural along with the outbreak related matches showed 100% query coverage, 100% sequence identity, and E value of 0.0 which indicates perfect conservation with no detectable amino acid differences.[6] Multiple other *Zaire ebolavirus* isolates showed 99-100% identity across the whole protein which further confirms the conservation of VP35 across strains.

This shows that VP35 sequence is highly conserved during persistence which suggests that viral persistence is unlikely to be happening by structural changes in VP35. Instead, persistence may need VP35 to be preserved as an antagonist of the innate immune system. That strong conservation that is seen in the different isolates supports that even small changes in VP35 may be bad for the viral fitness.

Additionally, the high sequence conservation suggests that selective pressure favors maintenance of VP35 functional domains. Also, mutations could disrupt important interactions that are crucial for survival in both acute infection and persistent virus because VP35 plays a key role in viral replication and immune evading. Therefore, it is anticipated that persistence is likely more of a function of host factors, immune modulation, and viral regulation than an adaptation through mutations in VP35.



Chain B, Polymerase cofactor VP35 [Ebola virus]
Sequence ID: **7YER_B** Length: 340 Number of Matches: 1
Range 1: 1 to 340

Score	Expect	Method	Identities	Positives	Gaps	Frame
711 bits(1835)	0.0()	Compositional matrix adjust.	340/340(100%)	340/340(100%)	0/340(0%)	

```
Query  1    MTTRTKGRGHTVATTQNDRMPGPELSGWISEQLMTGRIPVNDIFCDIENNPGLCYASQMQ  60
            MTTRTKGRGHTVATTQNDRMPGPELSGWISEQLMTGRIPVNDIFCDIENNPGLCYASQMQ
Sbjct  1    MTTRTKGRGHTVATTQNDRMPGPELSGWISEQLMTGRIPVNDIFCDIENNPGLCYASQMQ  60

Query  61   QTKPNPKMRNSQTQTDPICNHSFEEVVQTLASLATVVQQQTIASESLEQRITSLENGLKP  120
            QTKPNPKMRNSQTQTDPICNHSFEEVVQTLASLATVVQQQTIASESLEQRITSLENGLKP
Sbjct  61   QTKPNPKMRNSQTQTDPICNHSFEEVVQTLASLATVVQQQTIASESLEQRITSLENGLKP  120

Query  121  VYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAAATEAYWAEHGQPPPGPSLYEESA  180
            VYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAAATEAYWAEHGQPPPGPSLYEESA
Sbjct  121  VYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAAATEAYWAEHGQPPPGPSLYEESA  180

Query  181  IRGKIESRDETVPQSVREAFNNLDSTTSLTEENFGKPDISAKDLRNIMYDHLPGFGTAFH  240
            IRGKIESRDETVPQSVREAFNNLDSTTSLTEENFGKPDISAKDLRNIMYDHLPGFGTAFH
Sbjct  181  IRGKIESRDETVPQSVREAFNNLDSTTSLTEENFGKPDISAKDLRNIMYDHLPGFGTAFH  240

Query  241  QLVQVICKLGKDSNSLDIIHAEFQASLAEGDSPQCALIQITKRVPIFQDAAPPVIHIRSR  300
            QLVQVICKLGKDSNSLDIIHAEFQASLAEGDSPQCALIQITKRVPIFQDAAPPVIHIRSR
Sbjct  241  QLVQVICKLGKDSNSLDIIHAEFQASLAEGDSPQCALIQITKRVPIFQDAAPPVIHIRSR  300

Query  301  GDIPRACQKSLRPVPPSPKIDRGWVCVFQLQDGKTLGLKI  340
            GDIPRACQKSLRPVPPSPKIDRGWVCVFQLQDGKTLGLKI
Sbjct  301  GDIPRACQKSLRPVPPSPKIDRGWVCVFQLQDGKTLGLKI  340
```

Figure 1: BLAST alignment showing 100% sequence identity between persistence-derived VP35 and reference *Ebolavirus* VP35 (7YER_B), indicating strong conservation across infection states. Retrieved from reference 6

Comparative Sequence Alignment of Acute and Persistence VP35 Isolates:

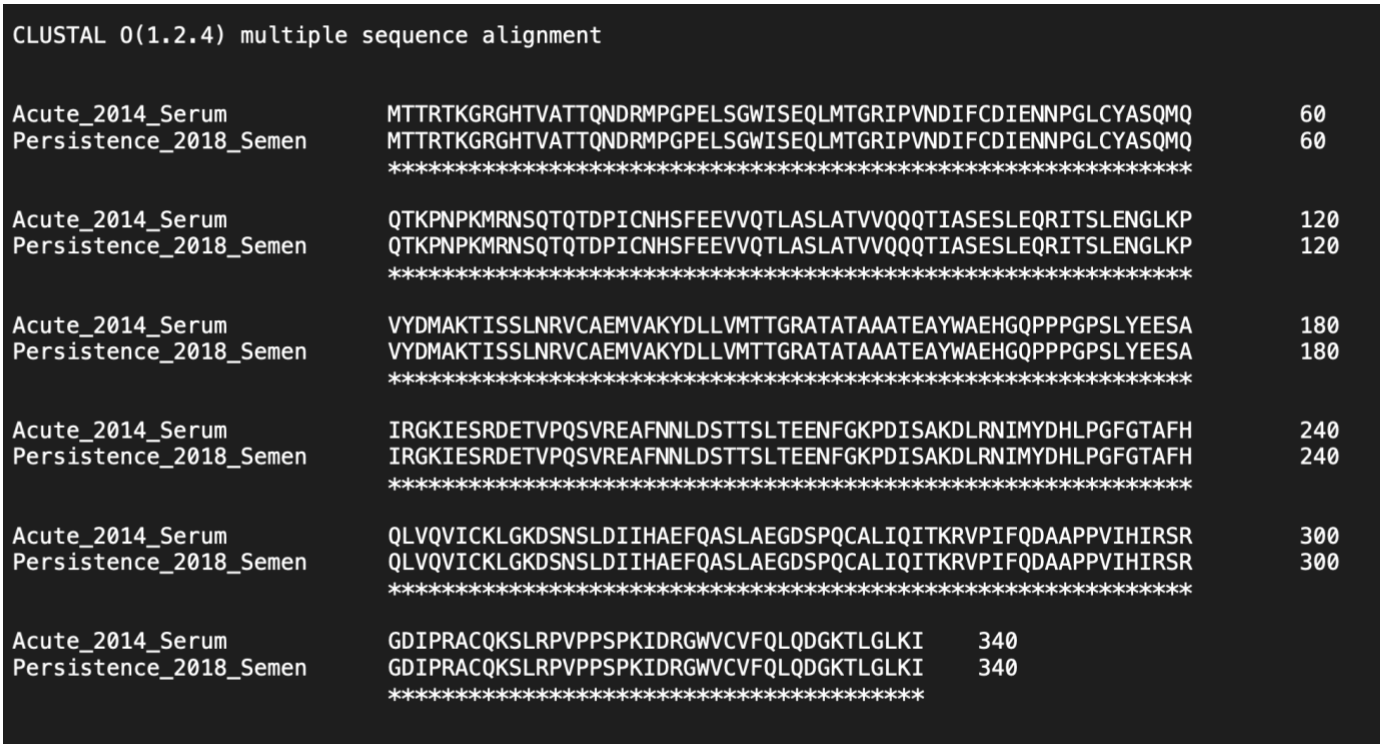
To further check if VP35 mutates during persistence, the amino acids sequence from acute infection and from a documented persistence case was compared. The acute infection sequence was obtained from a 2014 Makona lineage isolate which was collected from serum during active infection in Sierra Leone (Genebank protein AIE 11915.1, nucleotide accession

KM034562.1)[7] and the persistence sequence was also obtained from semen sample collected in 2015 from an *Ebolavirus* disease survivor in Sierra Leone (Genebank protein ARG43180.1 and nucleotide accession KY805810)[5] as reported in Whitmer et al 2018.

Multiple sequence alignment (Clustal Omega) was used to compare the two sequences, which represent the whole 340 amino acid VP35 protein.[8] The results showed that 100% alignment indicated 100% sequence conservation across all 340 amino acids. Between the two sequences, no insertions, deletions, substitutions, or other mutations were discovered (Fig. 2). Both the C terminal interferon inhibitory domain, which binds to dsRNA and stops the innate immune system signaling from starting, and the N terminal, which contains the LC8 motif, were included in the analysis.

This further proves the conservation of VP35 during persistence since there was no mutations detected during persistence. Although the persistence sample was collected around one year after the acute infection outbreak phase, the sequence of the VP35 remained unchanged. This conservation suggests that persistence does not require any modifications to VP35.

Having an identical VP35 protein sequence during persistence supports the biochemical possibility that VP35 continues to perform its role during persistence after recovery specially that the *Ebolavirus* mutates during persistence. When the full 19 kb virus sequences during active infection and during persistence were compared, there are many mutations which should show mutations in the VP35. Instead, complete conservation is seen and the VP35 protein is identical in both cases which suggests functional importance during persistence. Therefore, these findings support the idea that VP35 is included in the persistence model.



```
CLUSTAL O(1.2.4) multiple sequence alignment


Acute_2014_Serum            MTTRTKGRGHTVATTQNDRMPGPELSGWISEQLMTGRIPVNDIFCDIENNPGLCYASQMQ       60
Persistence_2018_Semen      MTTRTKGRGHTVATTQNDRMPGPELSGWISEQLMTGRIPVNDIFCDIENNPGLCYASQMQ       60
                            ************************************************************

Acute_2014_Serum            QTKPNPKMRNSQTQTDPICNHSFEEVVQTLASLATVVQQQTIASESLEQRITSLENGLKP       120
Persistence_2018_Semen      QTKPNPKMRNSQTQTDPICNHSFEEVVQTLASLATVVQQQTIASESLEQRITSLENGLKP       120
                            ************************************************************

Acute_2014_Serum            VYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAAATEAYWAEHGQPPPGPSLYEESA       180
Persistence_2018_Semen      VYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAAATEAYWAEHGQPPPGPSLYEESA       180
                            ************************************************************

Acute_2014_Serum            IRGKIESRDETVPQSVREAFNNLDSTTSLTEENFGKPDISAKDLRNIMYDHLPGFGTAFH       240
Persistence_2018_Semen      IRGKIESRDETVPQSVREAFNNLDSTTSLTEENFGKPDISAKDLRNIMYDHLPGFGTAFH       240
                            ************************************************************

Acute_2014_Serum            QLVQVICKLGKDSNSLDIIHAEFQASLAEGDSPQCALIQITKRVPIFQDAAPPVIHIRSR       300
Persistence_2018_Semen      QLVQVICKLGKDSNSLDIIHAEFQASLAEGDSPQCALIQITKRVPIFQDAAPPVIHIRSR       300
                            ************************************************************

Acute_2014_Serum            GDIPRACQKSLRPVPPSPKIDRGWVCVFQLQDGKTLGLKI    340
Persistence_2018_Semen      GDIPRACQKSLRPVPPSPKIDRGWVCVFQLQDGKTLGLKI    340
                            ****************************************
```

Figure 2: Clustal Omega multiple sequence alignment of full-length VP35 from acute infection isolate (2014 serum) and persistence isolate (2018 semen), showing complete sequence conservation across 340 amino acids. Retrieved from reference 8

Structural Conservation Analysis of VP35 using DALI:

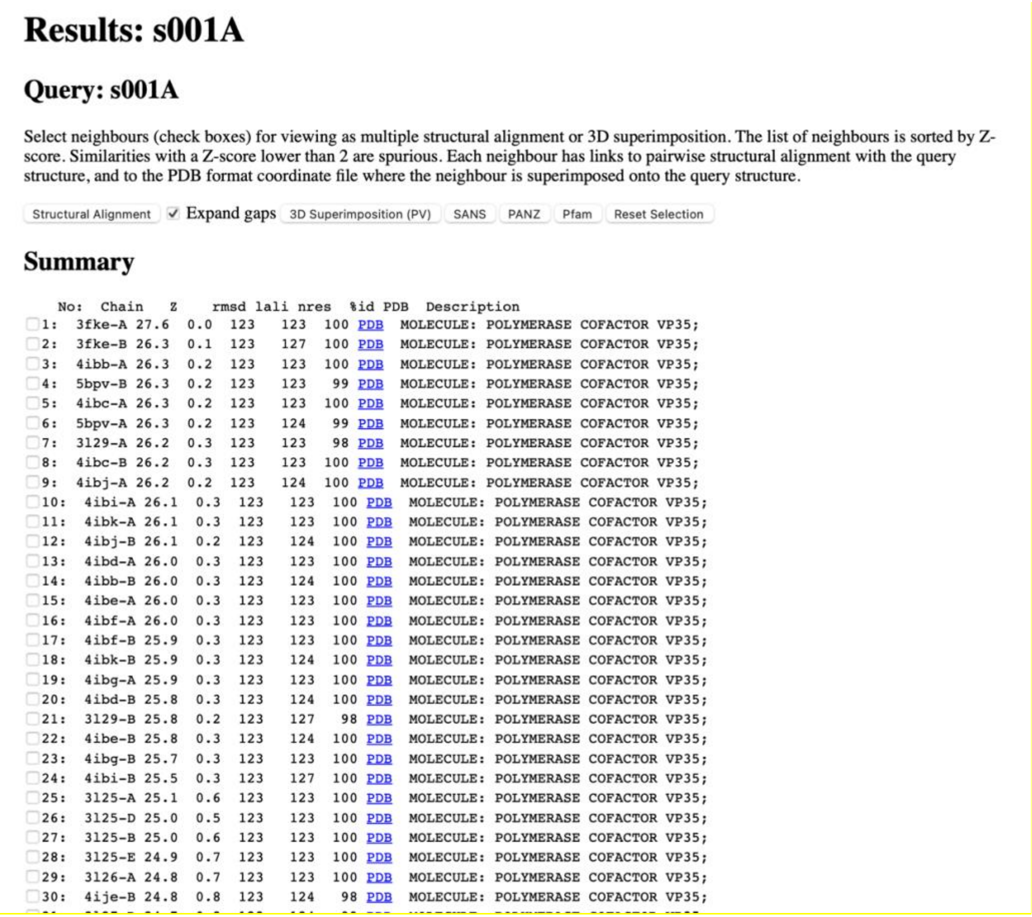
To determine how structurally conserved VP35 is, a structural analysis was performed on the structure of the *Ebolavirus* polymerase cofactor VP35 (PDB: 3FKE, chain A) using the DALI structure alignment server.[9] PyMol was used to delete chain B keeping only chain A to focus only on the VP35 domain without any other chains.[10] DALI ranked the structural neighbors based on the Z score which tells the similarity between the original structure and the neighbor.

Using VP35 (PDB: 3FKE, chain A) as the query, DALI identified top structural matches to other *Ebolavirus* VP35 structures, specifically the interferon inhibitory domain. The aligned structures showed very strong structural similiarity with Z score of 24 to 27 and RMSD value of 0 to 0.8 angstrom across 123 residues. A Z score greater than 20 indicates very strong structure similarity. This high Z proves near identical 3D structure along the VP35 crystal structures (Fig. 3). The analysis also showed a percent identity of 98% to 100% which confirms the conservation at both the sequence and the structure.

The DALI showed results like 4IBB, 4IBC, 4IBJ, and 4IBF which are crystal structures of the *Ebolavirus* VP35 interferon inhibitory domain (IID) bound to small molecule ligands that target its RNA binding surface. These structures will allow the comparison of whether ligand binding changes the overall protein fold, and the results show that all structures share almost identical backbone conformations with the query structure 3FKE which is shown in low RMSD values with only minor differences observed in flexible loop regions. Additionally, the beta sheet core and RNA binding interface remain structurally conserved across all the structures which suggests that ligand binding does not significantly disrupt the core structural features associated with VP35's role in immune evasion.

Also, the alignment covered about 123 residues or the major part of the folded region of the VP35 protein which means this region is the same for each crystal structure. The very low RMSD showed that the backbones are almost in the same positions in each of the structures which means that the fold of the protein does not change. The high structural conservation means that VP35 structure is very important to perform the role for the virus which is blocking the immune system of the host.

The high structural conservation supports the hypothesis that VP35 keeps a stable fold during both the infection and persistence, and since the structure is essential in protein interactions including binding to host factors such as dynein light chain, a host protein involved in intracellular transport, the conservation of the fold suggests that VP35 keeps it role during persistence as well. Both the sequence identity seen between infection and persistence and the DALI analysis provide strong evidence that persistence may require the VP35 role.



Results: s001A

Query: s001A

Select neighbours (check boxes) for viewing as multiple structural alignment or 3D superimposition. The list of neighbours is sorted by Z-score. Similarities with a Z-score lower than 2 are spurious. Each neighbour has links to pairwise structural alignment with the query structure, and to the PDB format coordinate file where the neighbour is superimposed onto the query structure.

Summary

```
    No:  Chain   Z    rmsd lali nres  %id PDB  Description
☐1:  3fke-A 27.6  0.0  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐2:  3fke-B 26.3  0.1  123    127  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐3:  4ibb-A 26.3  0.2  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐4:  5bpv-B 26.3  0.2  123    123   99 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐5:  4ibc-A 26.3  0.2  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐6:  5bpv-A 26.3  0.2  123    124   99 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐7:  3l29-A 26.2  0.3  123    123   98 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐8:  4ibc-B 26.2  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐9:  4ibj-A 26.2  0.2  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐10:  4ibi-A 26.1  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐11:  4ibk-A 26.1  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐12:  4ibj-B 26.1  0.2  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐13:  4ibd-A 26.0  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐14:  4ibb-B 26.0  0.3  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐15:  4ibe-A 26.0  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐16:  4ibf-A 26.0  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐17:  4ibf-B 25.9  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐18:  4ibk-B 25.9  0.3  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐19:  4ibg-A 25.9  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐20:  4ibd-B 25.8  0.3  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐21:  3l29-B 25.8  0.2  123    127   98 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐22:  4ibe-B 25.8  0.3  123    124  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐23:  4ibg-B 25.7  0.3  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐24:  4ibi-B 25.5  0.3  123    127  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐25:  3l25-A 25.1  0.6  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐26:  3l25-D 25.0  0.5  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐27:  3l25-B 25.0  0.6  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐28:  3l25-E 24.9  0.7  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐29:  3l26-A 24.8  0.7  123    123  100 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
☐30:  4ije-B 24.8  0.8  123    124   98 PDB  MOLECULE: POLYMERASE COFACTOR VP35;
```

Figure 3: DALI structural alignment results for *Ebolavirus* VP35, showing high structural conservation among homologous VP35 protein structures. Top matches exhibit Z-scores greater than 20 and low RMSD values, indicating near-identical three-dimensional folding across aligned residues. Retrieved from reference 9

Strctural Analysis of VP35 Interaction with Dynein Light Chain using PyMOL:

PyMOL was used to find the detailed structure of VP35 protein bound to human dynein light chain (LC8) which was obtained from PDB (7D35).[10] The PDB file was a VP35 chain bound to the dynein light chain. Molecular interface from PyMOL showed that the dynein light chain is made as a stable homodimeric structure composed of a central beta sheet surrounded by alpha helices. On the other end, the VP35 was seen forming a beta strand binding to the dynein binding groove which forms a beta sheet extension to the host protein.

The VP35 was seen to be aligned along the beta stands of dynein and creates the backbone hydrogen bonds which is consistent with the LC8 recognition motifs (Fig. 4). There were multiple hydrophobic and polar side chain residues of VP35 positioned deep in the dynein binding pocket with multiple interactions which provide stabilization. Specifically, the hydrophobic residues (Tyr65 and Val66) face the nonpolar regions of the groove which indicate that they are involved in packing interactions. The polar and charged residues (Glu69 and Thr67) are positioned to join in hydrogen bonding and electrostatic interactions with residues from the dynein pocket (Lys9, Asp12, Tyr75, and Phe73).

The interface has some degree of backbone beta sheet complementarity and specific side chain interactions which suggest that both the structural alignment and identity of the residues contribute to the stability of the complex. The peptide relies on the binding of dynein to maintain a structural fold which is consistent with motifs of LC8 binding proteins. That arrangement of this structure suggests that any disruption of key VP35 interface will result in great decrease in binding affinity and possibly hinder dynein mediated transport of the virus in the cell.

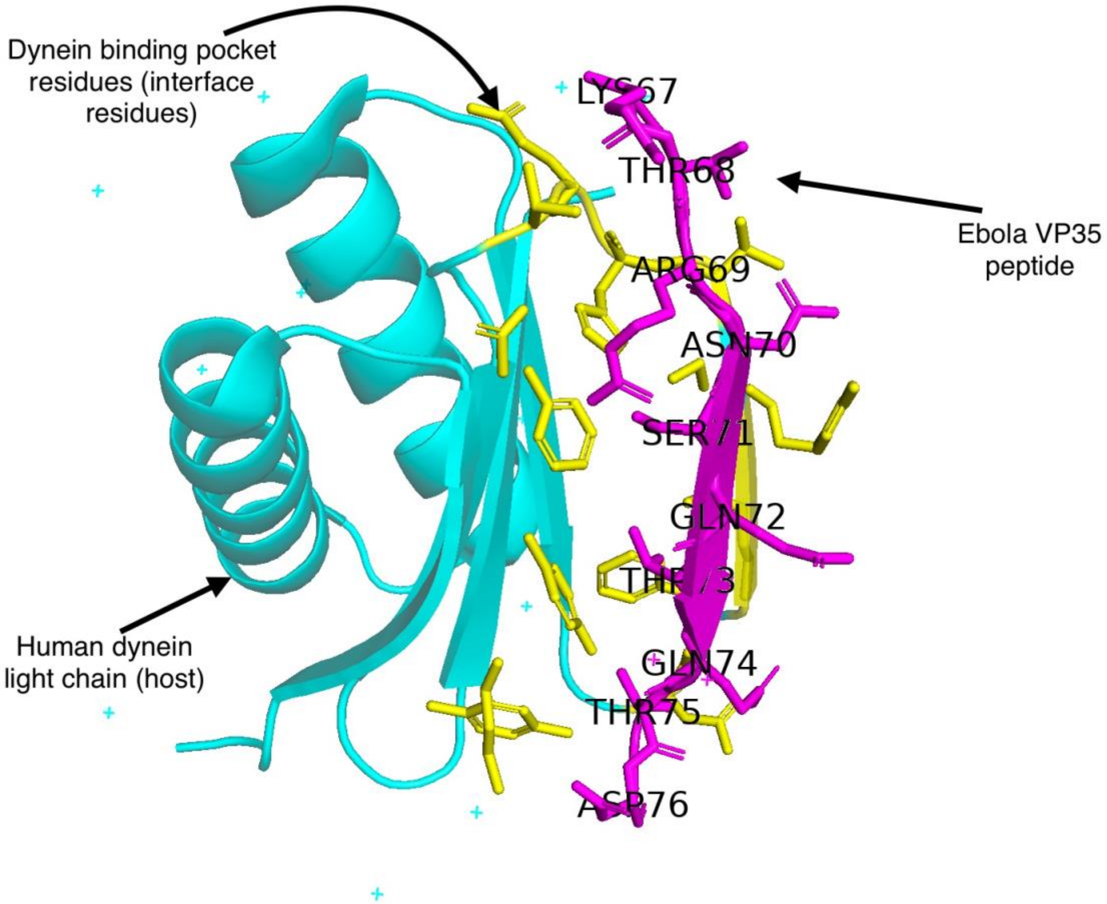


Figure 4: Structural interface between *Ebolavirus* VP35 peptide and human dynein light chain. The VP35 peptide residues (magenta) binds within a defined pocket on the dynein light chin (cyan). Dynein interface is shown in yellow, highlighting the binding pocket that mediates host viral interaction. Retrieved from reference 5

Energetic Contribution of VP35 Residues to Dynein Binding:

To determine the VP35 residues that play a critical role in binding to the dynein light chain (LC8), Bude Alanine Scan website was used to perform alanine mutation which evaluates how individual mutations affect binding stability through changes in ΔΔG.[11] This process identifies important residues by determining which mutations significantly destabilizes the interaction which suggests their role in maintaining binding. This process consists of conducting a computational alanine scan of all the original VP35 residues, substituting each for alanine one at a time, and then calculating the change in binding free energy (ΔΔG) for each mutation relative to the binding free energy for the complex.

Any mutation that produced a ΔΔG positive value has resulted in distortion of the original interaction and indicates that the original amino acid is energetically favorable in providing binding stability. Conversely, any mutation resulting in a ΔΔG negative or close to 0 has had minimal to no effect on providing binding stability.

Based on the results, the most significant contribution to the ΔΔG was seen in the B75 (Thr) to Ala mutation (11.37 kcal/mol) which indicates that this amino acid likely contributes greatly to the overall stability of the dynein binding interface (Table 1). The B70 (Asn) and B69 (Arg) mutations produced significant ΔΔG values of 7.12 kcal/mol and 6.76 kcal/mol suggesting that these amino acids are likely very important to the stability of the dynein binding interface. All three residues are in the center of the peptide and bind into the dynein site directly and it is reasonable to assume that all three are involved in the formation of essential hydrogen bonds and electrostatic interactions between the amino acids of the dynein binding cavity and those of VP35. Several other VP35 residues such as B76 (Asp), B74 (Gln), B73 (Thr), and B72 (Gln) also produced moderate destabilizing ΔΔG value (ΔΔG ~ 2.6-4.5 kcal/mol) indicating that, although these amino acids are not as critical as the center residues, they will likely also contribute to the binding stability of the interface.

The B68 (Thr) and B71 (Ser) mutations both produced slightly negative values for ΔΔG of -1.44 kcal/mol and -1.19 kcal/mol suggesting that these residues contribute little to the stability of the VP35 dynein interface. Therefore, computational alanine indicates that the central residues of the VP35 motif will be the most functionally important residues in stabilizing the position of the peptide in the dynein pocket. Additionally, the alanine scan suggests that mutations in these high ΔΔG residues could destabilize or prevent persistence.

Table 1: *Alanine scanning mutagenesis of VP35 interface residues showing the change in binding free energy (ΔΔG, kcal/mol) upon mutation to alanine. Positive ΔΔG values indicate destabilization of the VP35–dynein interaction, while negative values indicate minimal contribution to binding stability. Retrieved from reference 11*

Residue	ΔΔG (kcal/mol)
Lys67	0.61
Thr68	-1.44
Arg69	6.76
Asn70	7.12
Ser71	-1.19
Gln72	2.63
Thr73	3.27
Gln74	3.60
Thr75	11.37
Asp76	4.51

Functional Domain Conservation Analysis of VP35 During Persistence using InterPro:

InterPro is a very important database in bioinformatics that brings together multiple databases of protein identities which identifies the protein's family, domains, and functional sites by looking at the sequence of the protein.[12] The full sequence of VP35 from a persistence isolate (ARG43180.1, from semen of Ebola survivor) was analyzed in the InterPro database which identified the functional domains of VP35. The results showed that VP35 is classified a member of the family of VP35 proteins in the Filoviridae family, a polymerase cofactor, and an interferon antagonist.

The InterPro analysis showed that there is an evolutionary conserved interferon inhibitory domain (IID) present in the C terminal end of VP35 which overlaps with an area of VP35 that is known to bind to double stranded RNA and is associated with suppressing the host's innate immune signaling (Fig. 5). The IID is known to play a very important role for the inhibition of the virus RNA from being recognized by the pattern recognition receptors (PRR) such as RIG-I and MDA5. Also, the InterPro analysis showed areas that correspond to RNA binding functionality (V35-RBD-like domain) which is known to support the role of VP35 in interacting with viral RNA during replication. By mapping the conserved residues of VP35, clusters of highly conserved amino acids were present in both the IID and the RNA binding domains of VP35 which suggests that these functional sites remain conserved during persistence.

In summary, The InterPro analysis suggests that the key domains of VP35 specially the ones involved in immune evasion and RNA binding are highly conserved. Using the full sequence of VP35 during persistence showed that the functional domains remained unchanged during persistence which supports the possibility that VP35 may contribute to persistence.

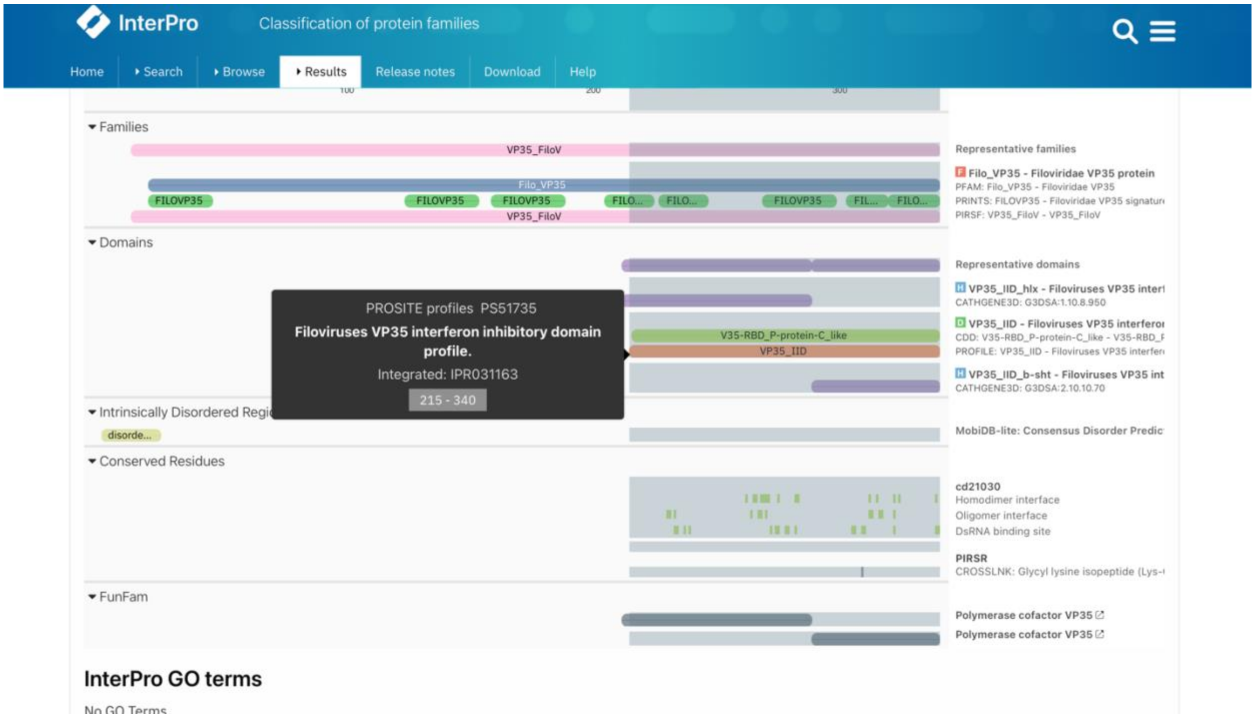


Figure 5: InterPro domain analysis of full-length VP35 from a persistence-derived isolate, showing classification within the Filoviridae VP35 protein family and identification of key functional domains. The conserved interferon inhibitory domain (IID) and RNA-binding regions are highlighted, indicating preservation of functional sites involved in immune evasion and viral replication. Retrieved from reference 12

Conclusion:

The purpose of this research was to see if it was biochemically feasible for VP35 (a protein that plays an essential role during Ebola infection) to play a role in Ebola persistence after recovery by evaluating the sequence conservation, structure stability, and functional interactions. The results from the BLAST and Clustal Sequence Alignment showed that VP35 is very conserved during acute infection and persistence which was seen in the 100% sequence conservation between the acute infection isolate and the persistence isolate. This high level of conservation of VP35 compared to the full genome of Ebola that mutates during persistence indicates that VP35 is likely under very strong selective pressure to retain its sequence, structure, and function.

The results from the structural analysis using DALI showed that the 3D structure of VP35 during persistence is nearly identical to the VP35 in the database which confirms the extremely close similarity in 3D structure of VP35 which is indicated by the high Z score and low RMSD value. This conservation of structure is very important because many of the important functions of a protein are defend by the shape that the protein takes to interact with the molecular partners. The PyMOL interface analysis showed that the interaction between VP35 and the dynein light chain (LC8) has a high degree of specificity and stability due to the presence of hydrophobic and electrostatic interactions located at the specific binding site. This interaction is very important for intracellular transport which supports VP35 beyond replication.

The alanine scan results showed the energetic contributions of individual residues within the interaction between VP35 and dynein light chain (LC8). The VP35 residues Thr75, Asn70, and Arg69 were all shown to destabilize the VP35 dynein interaction when mutated indicating that these three residues are critical for maintaining the stability of the binding site between VP35 and dynein. The presence of multiple amino acids with high ΔΔG suggests that that the interaction is highly optimized and sensitive to change which explains why VP35 stays conserved during persistence. Also, using the full length VP35 during persistence, the InterPro analysis confirmed that the major functional domains of VP35 (interferon inhibitory domain and RNA binding domain) remain conserved during persistence. These domains which are necessary for viral replication are responsible for stopping the activation of the immune system's interferon response.

These results provide computational evidence that VP35 is likely not changed or mutated during persistence, but actually conserved and retains conserved structural features associated with its functional roles which supports the hypothesis that VP35 may contribute to Ebola persistence, not through mutation, but through using its immune evasion functions. The VP35 ability to inhibit the signaling pathway of interferon signaling pathway provides a possible mechanistic explanation for why VP35 might be involved in persistence since persistence requires that the virus remains undetected by the immune system. However, although the results show that VP35 may be capable of supporting persistence, it also suggests that persistence is not driven by only VP35. It is very likely that other factors play a role such as other viral proteins, host immune responses, and tissue specific responses. Therefore, these conclusions are based on computational analysis which while powerful, but require experimental validation to confirm these findings.

Further studies should incorporate wet lab approaches to further analyze the role of VP35 in persistence. Structural studies such as x ray crystallography or cryo electron microscopy could provide higher resolution results about VP35 interactions. Also, mutagenesis in specific sites during persistence that would destabilize VP35 binding could provide more information about the necessity of VP35 during persistence. Comparing VP35 sequence during acute infection and during persistence in same patient could be helpful as well. Finally, other structural and binding studies can help further suggest other interactions during persistence.

References:

1. Ebola Disease Basics. Centers for Disease Control and Prevention. https://www.cdc.gov/ebola/about/index.html (accessed Feb 2026).

2. Durant, O.; Marzi, A. Ebola virus disease sequelae and viral persistence in animal models: Implications for the future. *PLOS Pathogens* **2024**, *20* (3) DOI: 10.1371/journal.ppat.1012065.

3. Cárdenas, W. B.; Loo, Y.-M.; Gale, M.; Hartman, A. L.; Kimberlin, C. R.; Martínez-Sobrido, L.; Saphire, E. O.; Basler, C. F. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by rig-I signaling. *Journal of Virology* **2006**, *80* (11), 5168–5178 DOI: 10.1128/jvi.02199-05.

4. Leung, D. W.; Ginder, N. D.; Fulton, D. B.; Nix, J.; Basler, C. F.; Honzatko, R. B.; Amarasinghe, G. K. Structure of the ebola VP35 interferon inhibitory domain. *Proceedings of the National Academy of Sciences* **2009**, *106* (2), 411–416 DOI: 10.1073/pnas.0807854106.

5. Whitmer, S. L. M.; Ladner, J. T.; Wiley, M. R.; Patel, K.; Dudas, G.; Rambaut, A.; Sahr, F.; Prieto, K.; Shepard, S. S.; Carmody, E.; et al. Active ebola virus replication and heterogeneous evolutionary rates in EVD survivors. *Cell Reports* **2018**, *22* (5), 1159–1168 DOI: 10.1016/j.celrep.2018.01.008.

6. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **1990**, *215* (3), 403–410 DOI: 10.1016/s0022-2836(05)80360-2.

7. Gire, S. K.; Goba, A.; Andersen, K. G.; Sealfon, R. S.; Park, D. J.; Kanneh, L.; Jalloh, S.; Momoh, M.; Fullah, M.; Dudas, G.; et al. Genomic surveillance elucidates ebola virus origin and transmission during the 2014 Outbreak. *Science***2014**, *345* (6202), 1369–1372 DOI: 10.1126/science.1259657.

8. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **2011**, *7* (1) DOI: 10.1038/msb.2011.75.

9. Holm, L. Using Dali for protein structure comparison. *Methods in Molecular Biology* **2020**, 29–42 DOI: 10.1007/978-1-0716-0270-6_3.

10. DeLano, W. L. The PyMOL Molecular Graphics System; Schrödinger, LLC: New York, NY.

11. Wood, C. W.; Ibarra, A. A.; Bartlett, G. J.; Wilson, A. J.; Woolfson, D. N.; Sessions, R. B. Balas: Fast, interactive and accessible computational alanine-scanning using Budealascan. *Bioinformatics* **2020**, *36* (9), 2917–2919 DOI: 10.1093/bioinformatics/btaa026.

12. Apweiler, R. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research* **2001**, *29* (1), 37–40 DOI: 10.1093/nar/29.1.37.

AI tools were not used to generate scientific content. All scientific content, analysis, and conclusions were reviewed and revised by the author.