Investigating Ebola virus pathogenicity using Molecular Dynamics

Morena Pappalardo[1§], Francesca Collu[2§], James Macpherson[2], Martin Michaelis[1], Franca Fraternali[2*], Mark N Wass[1*]

1. School of Biosciences, University of Kent, Kent, UK.
2. Randall Division of Cell and Molecular Biophysics King's College London

§These authors contributed equally
*Correspondence to Mark Wass m.n.wass@kent.ac.uk and Franca Fraternali f.fraternali@kcl.ac.uk

Keywords: Ebola, molecular dynamics, virology, protein structure, pathogenicity

Abstract

Background: Ebolaviruses have been known to cause deadly disease in humans for 40 years and have recently been demonstrated in West Africa to be able to cause large outbreaks. Four *Ebolavirus* species cause severe disease associated with high mortality in humans. Reston viruses are the only Ebolaviruses that do not cause disease in humans. Conserved amino acid changes in the Reston virus protein VP24 compared to VP24 of other Ebolaviruses have been suggested to alter VP24 binding to host cell karyopherins resulting in impaired inhibition of interferon signalling, which may explain the difference in human pathogenicity. Here we used protein structural analysis and molecular dynamics to further elucidate the interaction between VP24 and KPNA5.

Results: As a control experiment, we compared the interaction of wild-type and R137A-mutant (known to affect KPNA5 binding) Ebola virus VP24 with KPNA5. Results confirmed that the R137A mutation weakens direct VP24-KPNA5 binding and enables water molecules to penetrate at the interface. Similarly, Reston virus VP24 displayed a weaker interaction with KPNA5 than Ebola virus VP24, which is likely to reduce the ability of Reston virus VP24 to prevent host cell interferon signalling.

Conclusion: Our results provide novel molecular detail on the interaction of Reston virus VP24 and Ebola virus VP24 with human KPNA5. The results indicate a weaker interaction of Reston virus VP24 with KPNA5 than Ebola virus VP24, which is probably associated with a decreased ability to interfere with the host cell interferon response. Hence, our study provides further evidence that VP24 is a key player in determining Ebolavirus pathogenicity.

Background

The potential of Ebolaviruses to cause large outbreaks has been highlighted by the recent Ebola virus outbreak in West Africa [1] resulting in 28,657 confirmed cases and 11,325 deaths as of 8th May 2016 (www.who.int). To enable replication, viruses depend on mechanisms to antagonise the host cell interferon response. The Ebolavirus proteins that are known to be crucially involved in the suppression of interferon signalling are VP35 and VP24 [2-6]. VP35 prevents interferon signalling by binding and masking double stranded viral RNA. VP24 impairs the host interferon response by binding to the karyopherins α1 (KPNA1), α5 (KPNA5) and α6 (KPNA6) and the transcription factor STAT1 [3-5]. Karyopherins would normally bind STAT1 and transport it to the nucleus, a key step during interferon signalling. VP24 prevents this transport and the subsequent accumulation of STAT1 in the nucleus [2-6].

Of the five known *Ebolavirus* species, only Reston viruses are not pathogenic in humans [7]. A recent study identified genetic variants, that result in amino acid substitutions between the four human-pathogenic *Ebolavirus* species and Reston viruses that may explain the discrepancy in human pathogenicity [8]. Certain amino acid changes (T131S, M136L, Q139R – Ebola virus residue listed first and Reston virus residue second) in the VP24-karyopherin interface (Figure 1) were proposed to reduce the affinity of Reston virus VP24 to human karyopherins resulting in a (compared to other Ebolavirus VP24 proteins) reduced ability to inhibit interferon signalling.

In this study, we compared the interaction of Ebola virus and Reston virus VP24 with human KPNA5 using protein structural analysis and molecular dynamics simulations.

Results

Analysis of the effect of Ebola VP24 mutations on interaction with KPNA5.

The crystal structure of the Ebola virus VP24 complex with KPNA5 is available [3] as well as co-immunoprecipitation studies that have investigated the effect of mutations on the ability of Ebola virus VP24 to bind KPNA5. The experimental data has shown that the Ebola virus VP24 mutations R137A and the combination of R137A,T138A,Q139A strongly reduce VP24-KPNA5 binding. Combined F134A,M136A mutations resulted in an intermediate level of VP24-KPNA5 binding. Other single point mutations (including Q139A that we used as a control) had limited effect on VP24-KPNA5 binding [3].

Initially, we used mCSM [9] and FoldX [10] to predict the effect of the investigated mutations

on the stability of the Ebola virus VP24-KPNA5 complex (mCSM, FoldX) and the affinity of the proteins (mCSM) (see methods). mCSM predicted that both point mutations, R137A and Q139A, reduce the binding affinity and the stability of the complex, with the R137A mutation, however, having a greater effect (predicted ($\Delta\Delta$G -1.07 kcal/mol change in complex affinity) than Q139A (Table 1). Note that for mCSM negative $\Delta\Delta$G values indicate that the mutation is predicted to destabilise the interaction, while positive values stabilise the interaction. The FoldX predictions agree with mCSM for both point mutations. Additionally, FoldX was able to consider combinations of mutations simultaneously and predicted that both sets of mutations (F134A,M136A, and R137A,T138A,Q139A) reduce stability of the complex with a large reduction of more than -7 kcal/ml for the F134A, M136A combination (Table 1). These predictions are generally in agreement with the experimental observations that R137A and the two sets of multiple mutation reduce binding of Ebola VP24 with KPNA5 and that Q139A results in less pronounced effects [3].

Next, molecular dynamics simulations (MD) for both the wild type Ebola virus VP24 and mutated forms were performed. MD simulations were obtained for the wild type complexes and for the mutated Ebola virus VP24 forms: 1)R137A, 2)Q139A, 3)F134A,M136A and 4)R137A,T138A,Q139A. First the amount of positional fluctuation during the simulation for each residue in VP24 and KPNA5 was considered by calculating the Root Mean Squared Fluctuation (RMSF; Figure 2). This analysis showed that for all four mutation sets, a greater fluctuation variability is observed in KPNA5. Additionally, for all mutations there is increased fluctuation in VP24 in the interface and in the vicinity of the mutated residues (Figure 2). This fluctuation is much greater for the three sets of mutations that have an effect on the interaction between VP24 and KPNA5 and smallest for Q139A, which affects binding to a lesser extent (Figure 2).

Investigating the effects of VP24 R137A mutation on interaction with KPNA5

According to the experimental data from Xu et al [3] the single point mutation R137A is the only point mutation that largely removes binding of VP24 to KPNA5. Thus, we used this mutation to validate our system and to obtain additional molecular information on the interaction between the two proteins. The RMSF calculations (Figure 2) indicated the greatest movement compared to the wild type VP24 centred around residue 115 and also from residues 135 to 150.

To consider how this mutation affects the interaction between VP24 and KPNA5, we analysed the correlation of conformational changes within the complexes and investigated the solvation

properties at the interface. The solvation properties for the wild type and R137A-mutated VP24 complexes with human KPNA5 were calculated by estimating the water density on a grid of points constructed around the residues at the interface (see methods). This analysis gave a detailed description of how the solvation influences the binding of VP24 and KPNA5 (Figure 4).

By considering how frequently each grid point (represented as spheres in Figure 4) was visited by water molecules, we were able to identify solvation sites at the interface in a continuum range from "low density" grid points (short temporal permanence [ps] of water molecules, red spheres in Figure 4) to "high density" grid points (long temporal permanence [ns] of water molecules, blue spheres in Figure 4).

In the wild type Ebola virus complex, high density grid points were found close to the residues N185, H186, E203, P204 and D205 (Figure 4). This means that at the interface there are solvation sites where water molecules are trapped and contribute to stabilise the complex. In contrast, the mutant R137A VP24-KPNA5 complex at the interface is characterised only by low density grids points, meaning that water molecules do not strongly localise stably at this spot and consequently they are not likely to contribute to the stabilisation of the complex. The mutation induces a long-range destabilisation of the complex creating a cleft at the interface where water molecules can freely enter and exit (Figure 4B). This is an indication that the fit of the two protomers at the interface of the complex is not optimal. We interpret this conformational change as an early event of the detachment of the two proteins.

The correlation of conformational changes between VP24 and KPNA5 were considered next. Previous studies have found that the binding of a protein ligand to a protein receptor can result in correlated enthalpic backbone motions [11]. Consequently, a model of signal propagation built on the analysis of local motions was generated. These were extracted from the molecular dynamics simulations by encoding trajectories into sequences of 4-residue fragment states with the M32K25 structural alphabet ([12]; see materials and methods for a full description of the numerical procedure used) to estimate the propensity of the Ebola virus VP24 R137A mutant to remain bound to KPNA5. The local error of the structural alphabet fit for the three complexes is shown in figure S3.

The Ebola virus VP24-KPNA5 and Ebola virus VP24(R137A)-KPNA5 complexes were encoded with the M32K25 structural alphabet and their collective motions were analysed. We

assessed the quality of fragment encoding of the simulations by measuring the local fit root-mean square deviation of the fragments from the atomistic trajectories (Figure 3). We found that the fragment-based encoding produced simplified trajectories with a reasonable special accuracy (>1.0 Å) for each of the trajectories analysed. A statistical analysis of the global collective motions revealed that the Ebola virus VP24-KPNA5 complex displayed structurally correlated motions between distal fragments (Figure 3a). The molecular dynamics simulation of the mutant VP24(R137A)-KPNA5 interaction displayed fewer such collective motions between the two subunits in the dimer (Figure 3b). In support of this observation, a topological analysis of the three simulations revealed that the Ebola virus wild type VP24 complex displayed a denser network of inter-subunit correlations than the VP24(R137A) complex (Figure 3). Taken together, the physical model presented here would suggest that the wild type Ebola virus complex is molecularly optimised for dimerisation, whereas we predict that the mutant VP24(R137A) is unlikely to form a stable complex, which is in agreement with the experimental data from Xu et al. [3].

Comparison of Ebola virus and Reston virus VP24 interaction with KPNA5

Next, the complex formed between Reston virus VP24 and KPNA5 was investigated with the aim of identifying how the interaction between these two proteins may differ from the interaction between Ebola virus VP24 and KPNA5. Only an unbound structure of Reston VP24 was available. Hence, a model of Reston virus VP24 bound to KPNA5 was generated using the Ebola virus VP24-KPNA5 complex as a template (see Methods).

mCSM predicted that all of the conserved amino acid changes observed in Reston virus VP24 would reduce the stability and affinity of the complex, with the exception of M136L, where a small increase in affinity was predicted (Table 3). The other changes in stability were similar to the predicted change for R137A, which is known to reduce binding [3]. FoldX also predicted reduced stability for two (T131S, N132T) of these four point changes, with increased stability predicted for M136L and Q139R, although the $\Delta\Delta G$ for M136L is predicted to be very small (0.18 kcal/mol). FoldX also predicted a slightly less stable complex with all four amino acid changes present (Table 3). Overall these predictions suggest that together the amino changes at the Reston virus VP24-KPNA5 interface are likely to reduce the stability and affinity of the complex.

The interfaces of the Ebola virus VP24- and Reston virus VP24-KPNA5 complexes were compared (Table 2) at the beginning and end of the molecular dynamics simulation using PISA [9] and POPSCOMP [10]. Both complexes are stabilised by patterns of hydrogen bonds. In the X-ray structure of the Ebola virus VP24-KPNA5 complex, nine hydrogen bonds were present, while in the model of the Reston virus VP24-KPNA5 complex 11 hydrogen bonds were identified to form at the interface. In total three hydrogen bonds were equivalent in the two complexes for the energy minimised structures (i.e. at 0ns). Molecular dynamics simulations for Ebola or Reston virus VP24 in complex with KPNA5 were then performed for 600ns. We calculated the number of hydrogen bonds at the interface over the whole trajectory (600ns) for both complexes (Figure S4). From the probability distribution of the number of hydrogen bonds over the time, we observed that the Ebola virus VP24-KPNA5 complex is stabilised by 22 hydrogen bonds on average compared to 18 for the Reston virus VP24-KPNA5 complex (Figure S4). So during the simulation the complexes have a greater number of hydrogen bonds than the starting structures. The four fewer hydrogen bonds at the interface of the Reston virus VP24-KPNA5 would suggest that this complex is energetically weaker than the Ebola virus VP24-KPNA5 complex. Analysis of the simulations highlighted that in the Ebola virus VP24 complex there was a considerable hydrogen bonding network with KPNA5 (Figure 5B) involving VP24 residues 137-140. However, in the Reston virus VP24-KPNA5 complex there was only limited hydrogen bonding of VP24 R137 and T138 with KPNA5 D480 (Figure 5B). In addition, the Ebola virus VP24-KPNA5 complex contains a hydrogen bond between VP24 Q139 and KPNA5 E474. In contrast, VP24 R139 points away from KPNA5 and towards the solvent in the Reston virus VP24-KPNA5 throughout the simulation (Figure 5B).

Throughout the simulation, the RMSD of the main chain C-Alphas was stable for both complexes (Figure S1). The RMSD of the Reston virus VP24-KPNA5 model is greater than of the Ebola virus VP24-KPNA5 complex, (Figure S1). This could indicate a difference in the interaction between RESTV VP24 and KPNA5 and could also partly reflect that the simulation is based on a model rather than a solved structure.

For VP24 some minor differences in fluctuation (RMSF) were observed between the Reston virus and Ebola virus proteins. Two of these differences coincided with the interface site at residue 113 (Figure 2A). Residue 113 is located in an alpha helix at the interface. For KPNA5 there are larger differences in RMSF in four regions, three of which coincide with the complex interface (Figure 2A). The most pronounced difference is around residues 477-479 (a loop region between two alpha helices), where there is very little fluctuation of KPNA5 in the Ebola

virus VP24 complex (around 1 Å) but a peak fluctuation of 8 Å in the Reston virus VP24 complex. The greater fluctuation in KPNA5 suggests that the interaction with Reston virus VP24 differs from that with Ebola virus VP24, supporting the evidence that Reston virus VP24 and human KPNA5 are weaker binding partners than Ebola virus VP24 and human KPNA5.

Analysis of the secondary structure (using DSSP [13]; see methods) during the simulation revealed minor changes in the VP24 secondary structure occurring at the interface site (supplementary figure S2). The most important changes were found around residue 76 where there is a prevalence of turns in Ebola virus VP24 which is a coiled structure in Reston virus VP24. Residues 133 and 134 (Figure S2), as well as residue 146, which are proximal to the binding interface, lose their bend and beta bridge structures to become unstructured in the Reston virus complex. The largest changes in secondary structure were found in KPNA5, particularly in two regions between residues 365-375 and 385-395 (Figure S2). The second region, which is involved in binding VP24, loses alpha helical structure after 220 ns in the Reston virus complex and changes to a turn structure instead.

For the Reston virus VP24-KPNA5 complex, analysis of solvation at the interface identified high-density grid points (visited by permanent water molecules) close to the residues E203, P204, D205, D124 and R137 (Figure 4C). As in the case of the wild type Ebola virus VP24-KPNA5 complex, at the interface there are solvation sites where water molecules are trapped and contribute to the stabilisation of the complex. In both the wild type Ebola virus and Reston virus VP24-KPNA5 complexes high-density grid points were found close to the residues E203, P204 and D205. This means that these residues are important in enhancing the stability of both complexes while establishing favourable interactions with water molecules. These residues belong to a loop interacting with KPNA5 defining a cavity where the water molecules are trapped.

The presence in the Reston VP24-KPNA5 complex of high-density grid points, where water molecules are trapped, suggests that part of the interface between Reston virus VP24 and KPNA5 may be relatively stable compared to the R137A-mutated Ebola virus VP24-KPNA5 complex, where only low density grid points were observed, indicating an absence of water molecules that would stabilise the complex (Figure 4).

Correlation of conformational changes in the Reston virus VP24-KPNA5 complex was compared to the results for the wild type and R137A mutant Ebola virus VP24-KPNA5

complexes. No correlated motions were detectable in the simulation of the Reston virus complex variant of the VP24-KPNA5 complex (Figure 3c), whereas many were observed for the Ebola virus VP24-KPNA5 complex and fewer for the Ebola virus R137A VP24-KPNA5 complex (see above). Correlation of the conformational changes analysis supported this with the Ebola virus VP24-KPNA5 complex having a denser network of inter-subunit correlations than did the Reston virus VP24-KPNA5 complex, which reflects the results obtained for the R137A mutant Ebola virus VP24-KPNA5 complex (Figure 3). Taken together, these findings suggest that Reston virus VP24 forms a less stable complex with KPNA5 than Ebola virus VP24, in particular due to the similarity of the results obtained for Reston virus VP24 and R137A-mutated Ebola virus VP24, which is known not to bind human KPNA5 [3].

Discussion

This study investigated the dynamics of interaction of the Ebolavirus protein VP24 with human KPNA5. This interaction occurs during infection to prevent transport of STAT1 to the nucleus in order to inhibit interferon signaling [3-5]. Recent findings have identified positions that are differentially conserved between human-pathogenic Ebolaviruses and Reston virus, the only Ebolavirus that is not pathogenic to humans [8]. In particular three of these positions in the interface site between VP24 and human KPNA5 were suggested to affect binding of Reston virus VP24 to KPNA5 and to contribute to the lack of human pathogenicity of Reston viruses [8].

Experimental evidence showed the R137A mutation in Ebola virus VP24 to reduce VP24 binding to human KPNA5 [3]. Therefore, we compared the Ebola virus VP24-KPNA5 interaction with the R137A-mutant Ebola virus VP24-KPNA5 interaction to validate our system. In concert with the experimental findings [3], our analysis predicted that the R137A mutation reduces VP24-KPNA5 binding. Moreover, the correlations of conformational changes and solvation analysis provide novel molecular insights into how this mutation affects binding of the two proteins. They show that the mutation reduces the stability of the complex, enables the two proteins to move apart from each other, and for water to enter the interface.

The investigation of the interaction of Reston virus VP24 with human KPNA5 added further evidence that Reston VP24 is a weaker binding partner to human KPNA5 than Ebola virus VP24. This is most strongly suggested by the lack of correlated conformational movements between Reston virus VP24 and KPNA5. Some permanent water molecules are (in contrast to the R137A-mutated Ebola virus VP24-KPNA5 complex) observed in the interface of the

Ebola virus VP24- and Reston Virus VP24-KPNA5 complexes. This indicates some differences between the interaction of R137A-mutated Ebola virus VP24 and Reston virus VP24 with human KPNA5.

In conclusion, our results suggest that Reston virus VP24 forms a weaker complex with human KPNA5 than Ebola virus VP24. This weaker binding is anticipated to reduce (in comparison to Ebola virus VP24) the capacity of Reston virus VP24 to inhibit KPNA5-mediated STAT1 transport into the nucleus and to anatagonise interferon signalling in human cells. This reduced Reston virus VP24-KPNA5 complex stability is likely to contribute to the lack of human pathogenicity observed in Reston viruses. Hence, our findings contribute novel evidence indicating VP24 to be an important regulator of species-specific Ebolavirus pathogenicity, as previously suggested by the analysis of conserved differences between Reston viruses and human-pathogenic Ebolaviruses [8]. In addition, our findings provide novel mechanistic insights at the molecular level on the interaction of Ebolavirus VP24 with human KPNA5.

Methods

Modelling of a RESTV-VP24 KPNA5 complex

The Ebola and Reston virus VP24 sequences share 81.3% sequence identity and 96% similarity. The protein structures were aligned using Chimera [18] and a model for Reston VP24 in complex with human KPNA5 built using MODELLER 9.0 [19]. The Reston virus VP24 crystal structure (PDB 4D9O) and the Ebola virus VP24-KPNA5 complex (PDB 4U2X) were used as templates for the Reston virus complex model. 200 models were obtained and the one with the lowest DOPE score was selected.

Comparison of interfaces

PISA [9] and mCSM [9] were used to analyse the structural properties at the complexe interfaces, including solvent accessibility and binding affinity. FoldX [10] was used to predict the effects of the changes in Energy upon mutation, in terms of effects in protein stability. POPSCOMP [20] was used to determine the contribution of the individual residues to the hydrophilicity and hydrophobicity at the interface, according to their solvent accessible surface area (SASA), using default parameters. The residues were classified as being part of the core, support or rim regions of the interface according to the change in SASA (when the percentage of hydrophobicity was greater than 40 and difference in SASA was less than 10 $Å^2$ the residue was considered as core, otherwise it was rim).

Molecular Dynamics simulations

Molecular dynamics simulations were performed for the wild type forms of Ebola virus VP24 and Reston virus VP24 in complex with human KPNA5. Other simulations were performed on the Ebola virus VP24-KPNA5 complex with mutations introduced into VP24 where the effect on KPNA5 binding had been experimentally determined [3]. The mutations considered were: 1)R137A, 2)Q139A, 3)F134A,M136A and 4) R137A-Q139A.

Molecular dynamics simulations were performed using Gromacs 5.0.5 [21]. The initial structures were solvated with a cubic box of SPC water molecules, setting the minimal distance between the protein and the box boundaries to 12 Å. Ionisable residues were modeled in their standard protonation state at pH 7. The systems were then neutralised adding the appropriate number of counterions. The GROMOS 53a6 parameter set [22] was used. Periodic boundary conditions were imposed. All the bonds were frozen and a 2-fs time step was used. The Berendsen algorithm [23] was employed for temperature and pressure regulation, with coupling constants of 0.2 and 1 ps, respectively. The electrostatic interactions were calculated with the particle mesh Ewald method [24] with a 14 Å cutoff for the direct space sums, a 1.2 Å FFT grid spacing, and a 4-order interpolation polynomial for the reciprocal space sums. For Van der Waals interactions, a 14 Å cutoff was used. The neighbour list for non-covalent interactions were updated every 5 steps. The systems were first minimised with 1000 steps of steepest descent. Harmonic positional restraints with a force constant of 4.8 kcal/mol/Å−2 were imposed onto the protein heavy atoms and gradually reduced to 1.2 kcal/mol/Å−2 in 80 ps, while the temperature was increased from 200 to 300 K at constant volume. The system was then simulated at constant temperature (300 K) and pressure (1 bar) for 100 ps. After the removal of harmonic restraints, 2 ns of equilibration were run in NPT conditions. NPT production simulations were then run for 40 ns for each system. System coordinates were saved every 1 ps.

We selected the force field GROMOS 53a6 because is the latest version of the widely used force field [25] from the GROMOS family stemming from the most widely used GROMOS 43A1. These parameter sets are often selected in performance studies where force fields are assessed and compared [26-29], and where it is shown that it is still one of the best united-atom force fields available, providing in some cases even better results than some all-atom force fields. Additionally, it is known since a while that the 43A1 parameter set provides better alpha/beta relative stabilities than GROMOS parameter sets developed later [30].

600 ns MD trajectories were obtained for the Ebola virus VP24-KPNA5 complex and the model of the Reston virus VP24-KPNA5 complex. In our analysis, we omitted the first 280ns of the simulation, as this was the approximate time required for each of the systems to reach conformational equilibrium. 200 ns trajectories were obtained for R137A and F134A, M136A and 100 ns for all other simulations.

Molecular Dynamics Analysis

Trajectories were analysed using the GROMACS analysis tools, VMD tools and the Bio3D package for R [31,32]. For the analysis, standard Periodic Boundary Conditions were removed and Minimum Image Convention (MIC) were applied to all the trajectories. Rotational and translational movements were then deleted in order to perform the Principal Component Analysis. Secondary structure plots for trajectories were obtained using the DSSP [33] tool in gromacs. Root mean square deviation (RMSD) and fluctuation (RMSF) from the initial starting complex were obtained using Bio3D, as well as the PCA analysis and correlation plots.

Analysis of correlation of conformational changes

We used a fragment-based approach to simplify the local structure of each of the three molecular dynamics simulations. The structural alphabet M32K25 was used to encode fragments of four consecutive C^α atoms, so as to describe backbone conformations of the protein in a simplified manner [12]. Mutual information between two sets of fragments (i,j) was used to calculate the correlation in conformational changes using a numerical procedure outlined in [12]:

$$I^n = \frac{I(C_i; C_j) - \epsilon(C_i; C_j)}{H(C_i, C_j)}$$

where C_i and C_j are columns in the structural string alignment, $I(C_i;C_j)$ is the mutual information, $H(C_i;C_j)$ is the joint entropy and $\varepsilon(C_i;C_j)$ is the error.

Considering the fragment at positions C_i and C_j as a discrete distribution of fragments X and Y, we calculated the joint entropy $H(X,Y)$ as a pair of discrete random variables (X,Y) with a joint distribution $p(x,y)$ defined as

$$H(X;Y) = -\sum_{x\varepsilon X}\sum_{y\epsilon Y} p(x,y)logp(x,y)$$

where $p(x)$ and $p(y)$ are the associated marginal probabilities.

Analysis of solvation in the interaction surface

To perform the water density analysis, structures were simulated, as previously described, for 5 ns by MD with backbone restraints (1.2 kcal· mol−1· Å−2) to avoid any significant conformational changes of the protein during the simulation [31]. Water density maps were calculated at discrete points **r** defined by a 0.5-Å spaced rectangular grid around the solute. Each structure of the last 10 ns of the trajectory was superimposed to a reference to remove the overall roto-translational motion of the protein. The number density of the water oxygen atoms was then averaged at each grid point over the MD snapshots extracted every 0.1 ps and normalised by the bulk density evaluated in the 6-8 Å shell around the solute. The hydration sites were then identified as local maxima of the density map with g(**r**) > 1 and used to define the hydration score S_{hyd} as previously described [34]. Water density maps were calculated for representative conformations of the Ebola virus VP24-KPNA5 complex as extracted from the trajectory and compared to the solvent distribution for the representative conformations of Ebola virus R137A VP24-KPNA5 and Reston virus VP24-KPNA5 complex trajectories.

A good water model to use as a solvent in biomolecular simulations should be computationally efficient and at the same time reproduce accurately enough the properties of bulk water. Not to be underestimated, this model should be compatible with the force field used for the solute interactions. It is well known that simple effective pair potentials such as SPC, TIP3P and TIP4P are, in different measure, not able to accurately describe the entire range of water properties, nevertheless, they have all proven to successfully model water as a solvent in biomolecular simulations.

The general weakness demonstrated by these models is in the overestimation of the diffusion coefficient and the inaccurate description of the dielectric properties. On the other hand they are effective in the calculation of solute solvent energies and practical to use. In the past we have developed solvent density maps based on the SPC model [34,35]. We were particularly interested in the water-protein interactions due to the localization of water at the surface of the protein, therefore we used the same model in this application as a matter of consistency with the force field and our previous calculations.

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

Author Contributions

MNW and FF devised the research. MP, FC and JM performed experiments. MP, FC, JM, MM, FF and MNW analysed data. All authors wrote the manuscript.

References

1. Quaglio G, Goerens C, Putoto G, Rübig P, Lafaye P, Karapiperis T, et al. Ebola: lessons learned and future challenges for Europe. Lancet Infect Dis. Elsevier; 2016;16:259–63.

2. Cárdenas WB, Loo Y-M, Gale M, Hartman AL, Kimberlin CR, Martínez-Sobrido L, et al. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. J. Virol. American Society for Microbiology; 2006;80:5168–78.

3. Xu W, Edwards MR, Borek DM, Feagins AR, Mittal A, Alinger JB, et al. Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1. Cell Host Microbe. 2014;16:187–200.

4. Reid SP, Leung LW, Hartman AL, Martinez O, Shaw ML, Carbonnelle C, et al. Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. J. Virol. American Society for Microbiology; 2006;80:5156–67.

5. Reid SP, Valmas C, Martinez O, Sanchez FM, Basler CF. Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. J. Virol. American Society for Microbiology; 2007;81:13469–77.

6. Zhang APP, Abelson DM, Bornholdt ZA, Liu T, Woods VL, Saphire EO. The ebolavirus VP24 interferon antagonist: know your enemy. Virulence. 2012;3:440–5.

7. Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, Kawaoka Y, et al. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. Archives of Virology. Springer Vienna; 2010;155:2083–103.

8. Pappalardo M, Juliá M, Howard MJ, Rossman JS, Michaelis M, Wass MN. Conserved

differences in protein sequence determine the human pathogenicity of Ebolaviruses. Sci Rep. 2016;6:23743.

9. Pires DEV, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics. 2014;30:335–42.

10. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. Nucleic Acids Res. 2005;33:W382–8.

11. Fornili A, Pandini A, Lu H-C, Fraternali F. Specialized Dynamical Properties of Promiscuous Residues Revealed by Simulated Conformational Ensembles. J Chem Theory Comput. 2013;9:5127–47.

12. Pandini A, Fornili A, Fraternali F, Kleinjung J. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. FASEB J. Federation of American Societies for Experimental Biology; 2012;26:868–81.

13. Joosten RP, Beek te TAH, Krieger E, Hekkelman ML, Hooft RWW, Schneider R, et al. A series of PDB related databases for everyday needs. Nucleic Acids Res. 2011;39:D411–9.

14. Volchkov VE, Chepurnov AA, Volchkova VA, Ternovoj VA, Klenk HD. Molecular characterization of guinea pig-adapted variants of Ebola virus. Virology. 2000;277:147–55.

15. Ebihara H, Takada A, Kobasa D, Jones S, Neumann G, Theriault S, et al. Molecular determinants of Ebola virus virulence in mice. PLoS Pathog. Public Library of Science; 2006;2:e73.

16. Dowall SD, Matthews DA, Garcia-Dorival I, Taylor I, Kenny J, Hertz-Fowler C, et al. Elucidating variations in the nucleotide sequence of Ebola virus associated with increasing pathogenicity. Genome Biol. 2014;15:540.

17. Cross RW, Fenton KA, Geisbert JB, Mire CE, Geisbert TW. Modeling the Disease Course of Zaire ebolavirus Infection in the Outbred Guinea Pig. J. Infect. Dis. Oxford University Press; 2015;212 Suppl 2:S305–15.

18. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12.

19. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. Curr Protoc Bioinformatics. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2016;54:5.6.1–5.6.37.

20. Kleinjung J, Fraternali F. POPSCOMP: an automated interaction analysis of biomolecular complexes. Nucleic Acids Res. Oxford University Press; 2005;33:W342–6.

21. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1-2:19–25.

22. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem. John Wiley & Sons, Inc; 2004;25:1656–76.

23. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics. AIP Publishing; 1984;81:3684–90.

24. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. The Journal of Chemical Physics. AIP Publishing; 1995;103:8577–93.

25. Oostenbrink, C., Soares, T. A., van der Vegt, N. F. A., & van Gunsteren, W. F. (2005). Validation of the 53A6 GROMOS force field. *European Biophysics Journal : EBJ*, *34*(4), 273–284.

26. Cino, E.A., Choy, W.-Y., & Karttunen, M. (2012). Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, *8*(8), 2725–2740.

27. Cao, Z., Liu, L., Zhao, L., & Wang, J. (2011). Effects of different force fields and temperatures on the structural character of Abeta (12-28) peptide in aqueous solution. *International Journal of Molecular Sciences*, *12*(11), 8259–8274.

28. Lange, O. F., van der Spoel, D., & de Groot, B. L. (2010). Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophysical Journal*, *99*(2), 647–655.

29. Todorova, N., Bentvelzen, A., English, N. J., & Yarovsky, I. (2016). Electromagnetic-field effects on structure and dynamics of amyloidogenic peptides. *The Journal of Chemical Physics*, *144*(8), 085101.

30. Lin, Z., van Gunsteren, W. F., & Liu, H. (2011). Conformational state-specific free energy differences by one-step perturbation: protein secondary structure preferences of the GROMOS 43A1 and 53A6 force fields. *Journal of Computational Chemistry*, *32*(10), 2290–2297.

31. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics. Oxford University Press; 2006;22:2695–6.

32. Skjærven L, Yao X-Q, Scarabelli G, Grant BJ. Integrating protein structural dynamics and evolutionary analysis with Bio3D. BMC Bioinformatics. BioMed Central; 2014;15:399.

33. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983rd ed. 1983;22:2577–637.

34. Fornili A, Autore F, Chakroun N, Martinez P, Fraternali F. Protein-water interactions in MD simulations: POPS/POPSCOMP solvent accessibility analysis, solvation forces and hydration sites. Methods Mol. Biol. New York, NY: Springer New York; 2012;819:375–92.

35. De Simone, A., Spadaccini, R., Temussi, P. A., & Fraternali, F. Toward the understanding of MNEI sweetness from hydration map surfaces. *Biophysical Journal*; 2006;*90*:3052–61.

Figure Legends

Figure 1. Ebola virus VP24 complex with KPNA5. VP24 is coloured grey and KPNA5 is blue. Residues differentially conserved between Ebola and Reston viruses are shown in red (present in interface), orange (not in the interface site), they are and labelled with the Ebola virus amino acid, residue number followed by the Reston virus amino acid.

Figure 2: Root Mean Square Fluctuation (RMSF) for simulation of mutated Ebola virus VP24 with human Karyopherin RMSF for the mutants over the wild type complex. EBOV VP24 (on the left side of each graph) and KPNA5 (on the right) are shown. In black line the wild type complex RMSF is shown; in A), the red line shows R137A mutant, in B) the blue line shows the Q139A, in C) the magenta shows F134A-M136A mutant, in D) in cyan R137A-Q139A mutant are shown. In light blue the residues that are occurring at the interface are mapped under each curve.

 Figure 3: Correlation of conformational Changes in the VP24-KPNA5 molecular dynamics simulations. All-atom molecular dynamic simulations of a) Ebola VP24, b) Ebola VP24 R137A (shown in red) and d) Reston VP24 with human KPNA5 complex were encoded with the M32K25 structural alphabet and analysed for global correlated motions between distal backbone fragments. The top 10% most significant correlated motions were selected and are represented by yellow strings.

Figure 4: Solvation analysis of the VP24-KPNA5 complexes. In parts A, B and C the spheres represent the most visited grid points coloured from red to blue, with red being a low density grid point (short temporal permanence [ps] of water molecules) and blue a high density grid points (long temporal permanence [ns] of waters). A) High density grid points were found close to the residues N185, H186, E203, P204 and D205 in the Ebola virus VP24 with human KPNA5 complex. B) The Ebola virus mutant R137A VP24-KPNA5 complex at the interface is characterized only by low density grids points meaning that waters are not trapped and consequently do not contribute to the stabilization of the complex. The mutation induces a long-range destabilization of the complex determining the creation of a cleft at the interface where waters can freely enter and exit. C) For the Reston virus VP24 complex, the analysis of solvation at the interface identified high-density grid points close to the residues E203, P204, D205, D124 and R137.

Figure 5. Hydrogen bonding in the Ebola and Reston virus VP24 complexes with KPNA5. VP24 hydrogen bonding with KPNA5 in the region of residues R137-R140. Top two panels are for VP24 with the initial minimised structure (left) and the structure at the end of the simulation (right). Bottom panels show the equivalent for Reston VP24. In all images VP24 is coloured grey.

Table 1 Predicted effect on Ebola VP24-KPNA5 complex stability by mutations in Ebola VP24. Prediction from mCSM (stability and affinity) and FoldX (stability) were calculated for the mutated complexes.

Mutation	mCSM stability ($\Delta\Delta G$ - kcal/mol)	mCSM PP affinity ($\Delta\Delta G$ - kcal/mol)	FoldX stability ($\Delta\Delta G$ - kcal/mol)
R137A	-0.80	-1.07	-0.68
Q139A	-0.39	-0.24	-0.33
F134A,M136A	NA	NA	-7.3
R137A,T138A,Q139A	NA	NA	-1.02

Table 2: PISA and POSPCOMP Interface Analysis at 0 and 600 ns MD snapshots.

	Ebola VP24 –KPNA5 complex	Reston VP24 –KPNA5 complex
PISA results		
PISA results at 0 ns		
Interface Area (Å^2)	1099.7	1055.1
Solvation Free Energy (ΔiG, Kcal/M)	-8.5	-8.6
Hydrogen bonds	9	11
PISA results at 600 ns		
Interface Area (Å^2)	1119.2	1076
Solvatation Free Energy (ΔiG , Kcal/M)	-10	-9.1
Hydrogen bonds	7	11
POPSCOMP results after minimisation		
Hydrophobic difference (Å^2)	1042.28	1002.35
Hydrophilic difference (Å^2)	772.73	713.95
Total difference (Å^2)	1815.06	1716.43

Table 3: Predicted effect on Ebola virus VP24-KPNA5 complex stability by mutation of residues that are differentially conserved in Reston in Ebola VP24. Prediction from mCSM and FoldX are shown changes for the in the EBOV VP24 –KPNA5 complex.

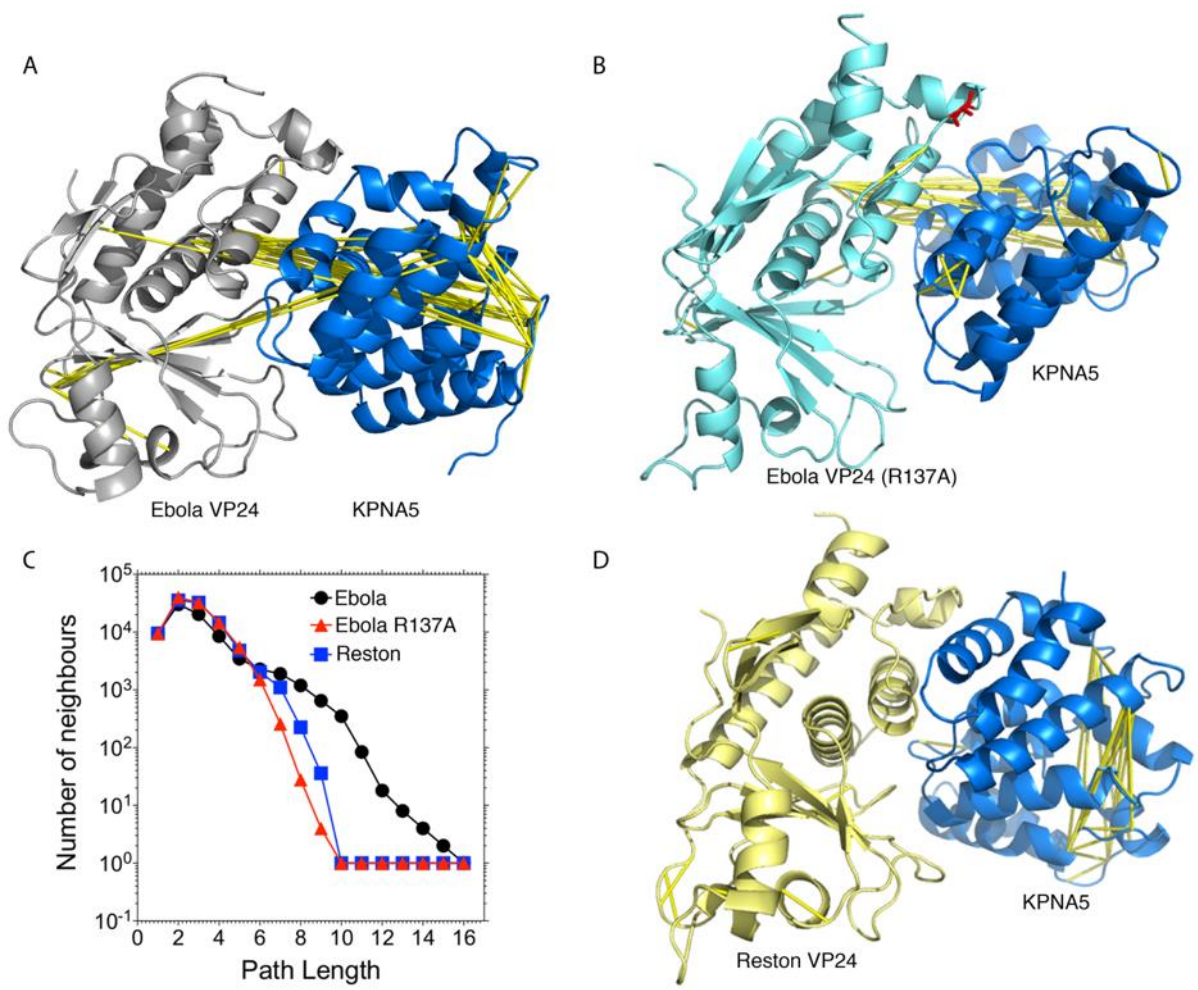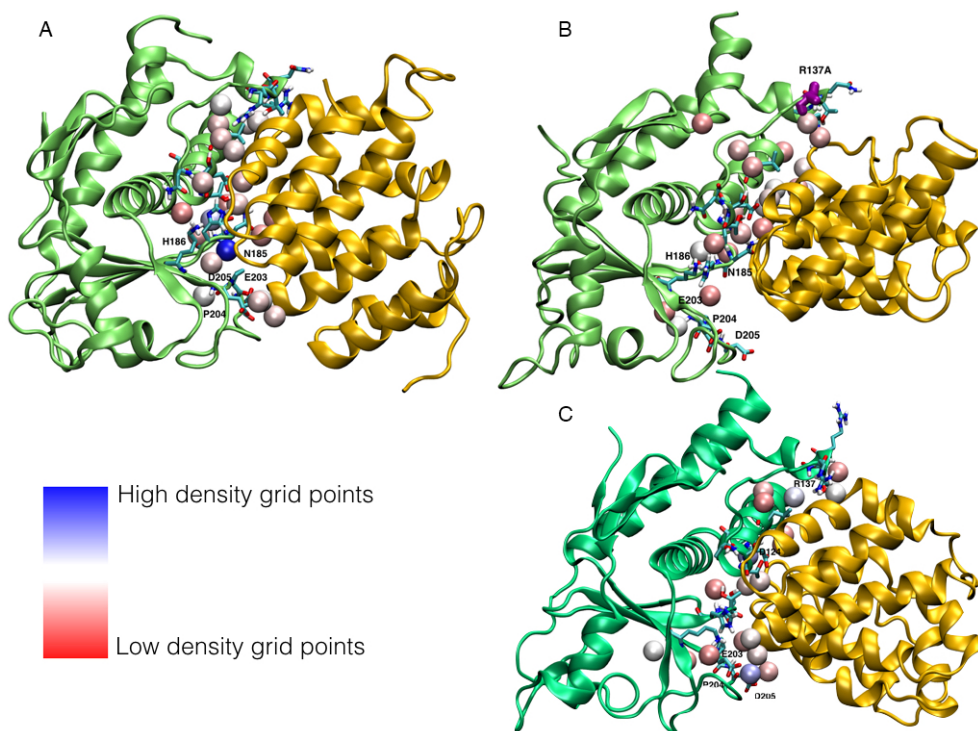| Mutation	mCSM stability (ΔΔG - Kcal/mol)	mCSM PP affinity (ΔΔG - Kcal/mol)	FoldX stability (ΔΔG - Kcal/mol)
T131S	-1.29	-0.32	-0.42
N132T	-0.62	-2.65	-1.22
M136L	-0.81	0.17	0.18
Q139R	-1.06	-0.99	1.59
T131S,M136L,Q139R	NA	NA	-0.3
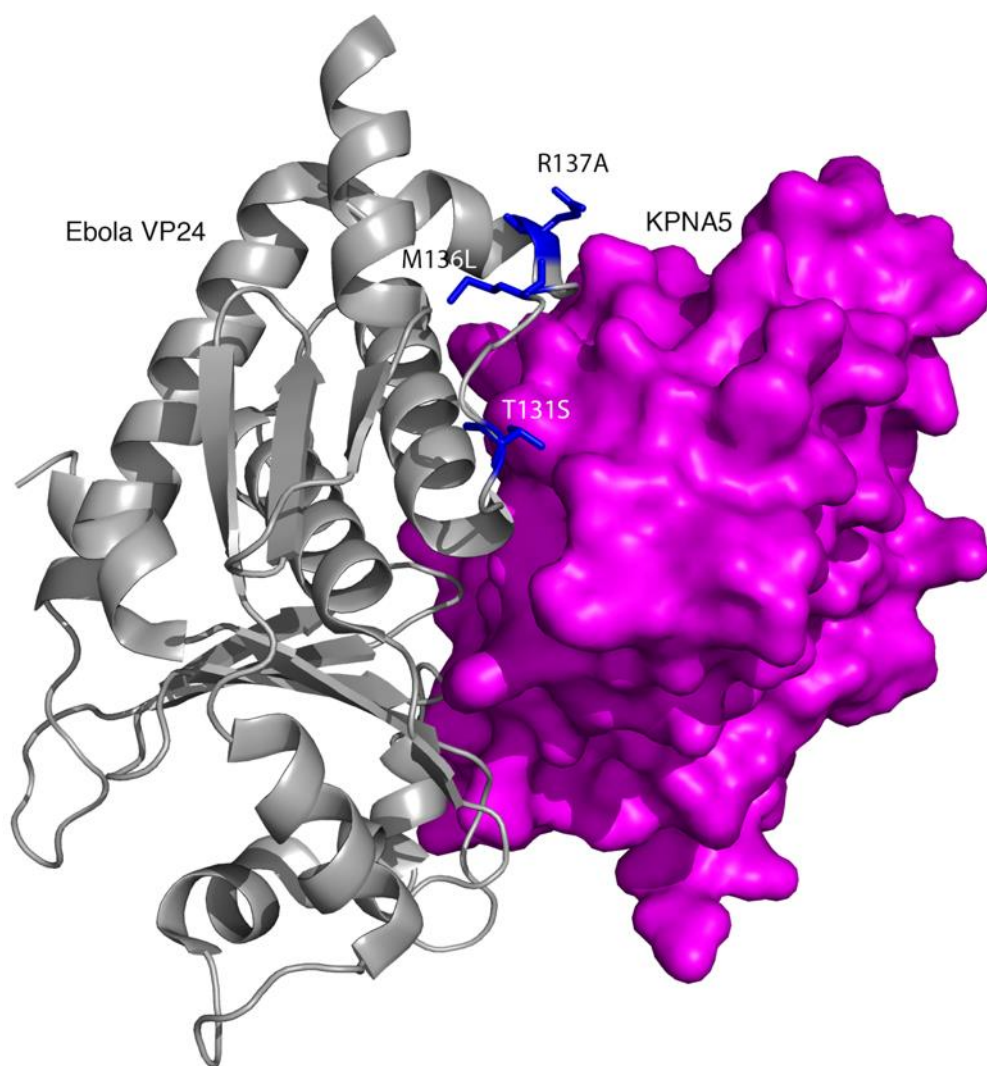
Figure 1

Figure 2

Figure 3

High density grid points

Low density grid points

Figure 4

Figure 5