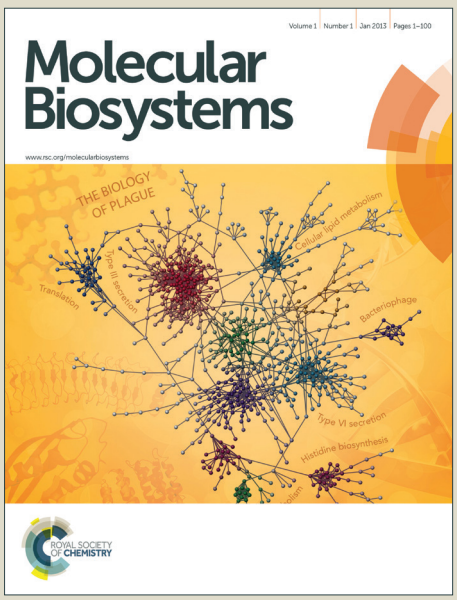
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

In Silico study of VP35 inhibitors: from Computational Alanine Scanning to Essential Dynamics

F. Dapiaggi[a], S. Pieraccini[a,b] and M. Sironi[a,b]

[a] Dipartimento di Chimica, Università degli studi di Milano, Via Golgi 19, 20133 Milano, Italy
[b] Istituto di Scienze e Tecnologie Molecolari (ISTM), CNR and INSTM UdR, Via Golgi 19, 20133 Milano, Italy

Abstract

In recent years Ebola virus has spread through several countries in Africa, recalling the need to develop new treatments for this disease and boosting a new research effort on this subject. Ebola virus Viral Protein 35 (VP35) carries out multiple functions necessary for virus replication and infection, in particular interfering with the (IFN)-α/β signaling. Recently, VP35 has been crystallized in complex with small organic molecules able to inhibit its interaction with Viral Nucleoprotein, thus reducing Ebola infection of cultured cells. In this work, starting from these structures, we carry out a computational study aimed at the investigation of the energetic and dynamical aspects of the interaction between VP35 and its ligands at the atomic level. Molecular dynamics simulations, Computational Alanine Scanning, Root Mean Square Fluctuations bootstrap analysis and Essential Dynamics analysis were performed. Our results expand the experimental ones obtained in previous works, adding information about the interactions landscape with the identification of a set of new hot-spots residues exerting a critical function in the protein-ligand interaction. Moreover we characterized the dynamics of the complexes, showing that the presence of ligands modifies the overall protein dynamics as well as the behavior of particular protein segments.

Introduction

Ebola virus (EBOV) and Marburg virus (MARV) are the causative agents of hemorrhagic fevers with a very high human fatality rate, near 90%.[1] Recently EBOV has spread in several countries in Africa and few sporadic cases were registered also in Europe and in the US. The lack of effective therapies or vaccines, combined with this high mortality rate, makes urgent the need to develop antiviral drugs against EBOV.

The virulence and high lethality of this virus are due to different factors, in particular to its ability to inhibit both the innate immune response in the early stages of infections and the subsequent adaptive specific immune responses of the host organism.[2-3] This is obtained by different strategies, in particular by the suppression of interferon (IFN)-α/β production and inhibition of interferon induced antiviral activity.[4-5]

Ebola's genome codes for seven proteins. Only one of them , the viral large (L) RNA-dependent RNA polymerase protein, has an enzymatic activity, while all the remaining EBOV proteins exert their function through protein-protein interactions, among them or with human proteins.

In this work we focus our attention on Viral Protein 35 (VP35). This protein carries out multiple functions, including the inhibition of RNA-dependent protein kinase (PKR)[6-7], suppression of RNA silencing[8] and the inhibition of (IFN)-α/β signaling.[9] This makes VP35 a potential therapeutic target. VP35 consists of a N-terminal coiled-coil domain, required for its oligomerization and of a C-terminal domain, called interferon inhibitory domain (IID), required for the interaction with the viral nucleoprotein (NP) and for interferon inhibition. The high resolution structure of the C-terminal domain has been recently solved. It is made up by two different subdomains (Fig. 1): a N-terminal α-helical subdomain and a C-terminal β-sheet subdomain.[10-11]

In a recent study Brown *et al*.[10] determined crystallographic structures of VP35 in complex with nine small molecules able to inhibit the protein-protein interaction between VP35 and NP.

In this work we study at the atomic level the interactions between these inhibitors and VP35 by performing molecular dynamics simulation and Computational Alanine Scanning on these systems, in order to identify the key interactions for the binding of ligands, thus providing hints for further optimization of active molecules targeting VP35. Then, to better characterize and

distinguish the dynamic behavior of the apo-protein from the ligand-protein complexes we measured residues RMSF and performed an Essential Dynamics analysis. We finally compared the dynamic behavior of apo-forms of EBOV and MARV VP35.

Materials and Methods

Molecular Dynamics

The ligand-protein complexes and the apo-protein structures of EBOV and MARV VP35 were obtained from the Protein Data Bank (PDB codes: 4IBB, 4IBC, 4IBD, 4IBE, 4IBF, 4IBG, 4IBI, 4IBJ, 4IBK,[10] 3FKE[11], 4GH9[12]). Molecular dynamics simulations were performed with the GROMACS 4.5.3[13] package using explicit solvent and periodic boundary conditions. The AMBER99SB-ILDN[14] force field was used for the protein and the Generalized Amber Force Field[15] for ligands. Each complex was solvated with TIP3P[16] waters and neutralized with two chloride ions. The LINCS algorithm[17] was employed to constrain all bonds involving hydrogen to their equilibrium length, allowing a time step of 2 fs. The systems were submitted to 10000 steps of geometry optimization with the steepest descent method. Then they were equilibrated for 200 ps in NVT condition (T=300 K) and subsequently for 200 ps in NPT conditions, in order to equilibrate systems density. Then a 25 ns molecular dynamics was performed for every system in NPT conditions (1 bar, 300 K). Temperature and pressure were kept constant to their reference values using the velocity rescale algorithm[18] and the Berendsen barostat respectively.[19] A 14 Å cutoff was applied for non-bonded interactions and the Particles Mesh Ewald algorith[20] was employed to calculate long range electrostatic interactions.

Computational Alanine Scanning

To perform Computational Alanine Scanning, 250 snapshots were extracted from the last 5 ns of each dynamics (one snapshot every 20 ps). Residues at the ligand-protein interface were selected using Naccess.[21] For each residue at the interface all sidechain atoms beyond C_β were removed and the missing hydrogen was added, so obtaining an alanine sidechain. ΔG of binding was calculated using the MM/PBSA approach.[22] In this protocol it is implicitly assumed that point mutations in the protein do not significantly affect the conformation of the mutated protein. The validity of this assumption in Computational Alanine Scanning has been widely confirmed in literature when it has been applied both to protein-protein and to protein-small molecule complexes.[23-26] The Poisson-Boltzmann equation was solved with APBS[27] using a relative dielectric constant of 80 for the region around the protein and of 2 for the protein interior. The entropic contribution to ΔG is supposed constant in the mutated and wild type structure considering their similarity , so it has not been calculated, as discussed by Kollman *et al.*[28]

Bootstrap root mean-square fluctuations (RMSF) analysis

For this analysis we applied the protocol described in the work by Mitra *et al.*.[29] We first determined the number of independent conformations in our MD trajectories by calculating the autocorrelation function of RMSD of the structures. We found a correlation time of 2.5 ns, corresponding to 10 independent structures in each of our 25 ns trajectories. In order to calculate the RMSF we selected 10 random points with replacement, from every dynamics and repeat it 200 times, giving us a mean RMSF value and a standard error of the mean for each residue in the protein. Using these values we applied a student's t-test to the data and identified contiguous regions that showed significant differences at a significance level of $p < 0.0005$.

Essential Dynamics

Principal component analysis or Essential Dynamics[30] (ED) extracts the correlated motions of a protein to understand which are the most important ones. First of all, the 3Nx3N covariance matrix of positional fluctuations of C_α was calculated:

$$\mathbf{C} = < (\mathbf{x}(t) - < \boldsymbol{x} >)(\mathbf{x}(t) - < \boldsymbol{x} >)^T >$$

where **x**(t) are the atomic coordinates in function of time and brakets denote an ensemble average. Then covariance matrix was diagonalized by an orthogonal coordinate transformation **T**:

$$\mathbf{C} = \mathbf{T}\Lambda\mathbf{T}^T$$

Λ is the diagonal matrix containing eigenvalues and every column of **T** is a eigenvector of **C**. Finally the dynamics of the complexes were projected on the first eigenvectors in order to analyze the motion. The covariance matrix and its diagonalization were calculated with g_covar utility of GROMACS 4.5.3, the analysis of the eigenvalues and eigenvectors with g_anaeig. The visual inspection of the projection on the first 3 eigenvectors was carried out using the VMD software.[31]

Results and Discussion

Computational Alanine Scanning

A molecular dynamics of 25 ns was performed for the apo-protein and for every ligand-protein complex. Computational Alanine Scanning (CAS) on the residues at the protein-ligand interface was carried out on the last 5 ns of the molecular dynamics trajectories, when the system was fully equilibrated in all cases, as shown by RMSD plots (see Supplementary Fig. S2-S10). In general, once equilibrium has been reached, trajectories 5 ns long guarantees enough sampling for accurate binding free energy calculation with the MM/PBSA approach and consequently for computational alanine scanning purposes.[32-33] All the complexes exhibit a negative protein-ligand ΔG of binding during the simulations (Supplementary Table 1). The interface is defined as the ensemble of amino acids whose solvent exposed surface area has a non-zero variation upon complex formation. Ligand molecules share the same binding site and 18 residues in total were found to be at the interface for each molecule. CAS results are summarized in Table 1. Residues whose mutation led to a variation in ligand binding free energy (ΔΔG) greater than 2 kcal mol^{-1} are defined hot-spots.

Residue numbering and domain names are taken from Leung *et al.*.[11] In Fig.1 the two subdomains of VP35 are shown. The α-helical subdomain consists of residues Ala221-Arg283, the β-sheet subdomain is formed by residues Val294-Ile340. The loop Val284-Pro293 connects the two subdomains. The binding site is the same for all the ligands and it is located between the α-helical and the β-sheet subdomains.

Six aminoacids are found to be hot spots in all the simulated complexes, namely Lys222, Arg225, Gln244, Lys248, Lys251 and Ile295 (Fig. 2). Brown *et al.*[10] performed experimental mutation of Lys248 or Ile295 into alanine, showing that these two residues are of paramount importance for ligands binding, as their mutation resulted in a near complete loss of binding affinity for all the tested ligands. Our computational results mirror the available experimental ones, furthermore led to the identification of four new hot-spots residues exerting a critical function in the protein-ligand interaction.

All the molecules considered in the present study share an almost equal orientation within the binding pocket, with the common carboxylic group oriented toward a region of the protein containing a cluster of positive charged residues, such as Lys222, Lys251, Lys248 and Arg225 which are all hot spots, the most prominent two being Lys248 and Lys251. Lys251 forms a strong salt bridge with the carboxylic group of the ligands, which is observed in all the simulations. The loss of this interaction upon mutation into alanine explains the high value of ΔΔG of the residue in all the complexes. In some cases, such as the GA017, GA246 and VPL29 complexes, the average distance between the carboxylic carbon and the charged lysine nitrogen is below 4 Å, the interaction being almost continuously present during the MD run. In other cases, for example the VPL27 complex, the average distance is around 5 Å, resulting in a lower ΔΔG. Lys248, although bearing a positive charge, also forms hydrophobic interactions through its aliphatic chain with the B and D rings of the ligands (Fig. 3). A coulombic interaction between opposite charges of Lys 248 and the carboxylic group of the ligands is likely to be also present. Analysing the geometries of the complexes during the simulations and measuring the average distance between the carboxylic carbon and the charged lysine nitrogen , it is clear that the contribution of hydrophobic interaction is constant for all the complexes, while the different values of Lys248 ΔΔG are modulated by electrostatic interactions during the dynamics. For VPL29 and VPL57, showing the higher ΔΔG value, the mean distance is around 5.5 Å, instead for VPL48 and GA246 this value is over 7 Å.

Lys222 and Arg225 don't form stable salt bridges during the dynamics, but they are affected by the proximity of ligand's carboxylic group. Their ΔΔG values are lower than those of the aforementioned positive charged residues.

Ile295 interact with the D and C aromatic rings of the ligands; these interactions are mainly hydrophobic. Interestingly, Asp252 exhibit a negative ΔΔG in all the complexes. This is probably due to its proximity to the negative charged carboxylic group of the ligands. Substituting this residue with an alanine results in the loss of this disadvantaged interaction. Asp302 exhibits the same behavior, albeit to a lower extent (see Table 1).

Root mean-square fluctuations bootstrap analysis

To further characterize the dynamic behavior of EBOV VP35 and how it is affected by ligand binding, C_α root mean-square fluctuations were calculated for the apo-form and for each protein-ligand complex.

As we can notice in Fig. 4, there are five protein segments exhibiting a fluctuation larger than average, corresponding to residues His231-Phe235, Ile246-Ser253, Ser266-Ser272, Pro293-Ile297 and Phe328-Leu336. Segments comprising residues His231-Phe235 and Ser266-Ser272 correspond to loop regions connecting α-helices, which have a intrinsically larger mobility than neighboring

residues, so we focused our analysis on Ile246-Ser253, Pro293-Ile297 and Phe328-Leu336 segments (Fig. 5), which all show significant differences in RMSF between the apo-protein and the complexes, according to student's t-test. Fig. 4 shows that residues ranging from Pro293 to Ile297, belonging to β-sheet subdomain exhibit a higher RMSF value in the apo-form than in all the complexes. It is also evident a similar behavior in residues ranging from Ile246 to Ser253 belonging to α-helical subdomain. The presence of any ligand reduces the fluctuation of these two regions which compose the binding site, and they are thus less flexible upon ligand binding.

Moreover, RMSF exhibits a sharp peak corresponding to the region between residues Phe328 and Leu336. Here RMSF reaches values around 0.2 nm for the apo-form and even greater for VPL60, VPL57 and VPL58 complexes. These ligands seem to exalt the flexibility of these residues. Finally, root mean square fluctuation analysis was performed on MARV VP35. On the whole, more flexible loops correspond to those observed in EBOV protein (Supplementary Fig. S12), as expected due to the high homology and great structural similarity between the two proteins.

Essential Dynamics

In order to better understand the results arising from RMSF analysis we performed essential dynamics (ED) calculations on the apo-protein and on all the simulated complexes. The MD trajectories were processed by principal components analysis considering the protein C_α for the construction of the covariance matrix. For each system, the projection of the trajectory on the three eigenvectors corresponding to the greater eigenvalues were visualized. In all the systems the first three eigenvectors represent at least 30% of the global motion.

The ED analysis of the apo-form of VP35 highlight two kinds of motions. The projection along the first two eigenvectors describes a closing motion of the β-sheet subdomain on the α-helical subdomain. In particular loops Pro293-Ile297 and Phe328-Leu336, both belonging to the β-sheet subdomain, appears to bend towards residues Ile246-Ser253 of the α-helical subdomain (Fig. 6). The projection on the 3rd eigenvector shows the motion of the Phe328-Leu336 loop alone. This results strictly mirror the observations from the RMSF analysis, where the aforementioned groups exhibited the higher mobility. For the apo-form the first three eigenvectors describe more than 50% of the collective motion of the protein.

The presence of some ligands, such as GA017, GA246 and VPL29 hinders the closing motion observed in the apo-form; the projection on the first eigenvector of GA017 complex is the same of the apo-form, but the amplitude of the motion is smaller. Also in this case we confirm our previous observation from RMSF analysis, as RMSF values for residues Ile246-Ser253 and Pro293-Ile297 of the complexes were lower than for the apo-protein.

Interestingly, RMSF analysis suggested that ligands VPL57, VPL58 and VPL60 could exalt the motion of residues between Phe328 and Leu336. The projections on the first three eigenvectors for VPL60 complex trajectory show an intense movement of these residues. In this case the characteristic closing motion observed in the apo-form was not detected , while it was significantly reduced in VPL57 and VPL58 complexes. For VPL60 complex the first three eigenvectors describe about 48% of the global motion. For VPL57 complex the projections on the 1st and 2nd eigenvector show a very large motion of the loop Phe328-Leu336 (Fig. 7) while the projection on the 3rd eigenvector describes the closing motion of the β-sheet subdomain towards the α-helical subdomain.

Finally, VPL58 has a similar behavior, the projection of the trajectory on the 1st eigenvector resulting in a large motion of the Phe328-Leu336 loop and the projections on the 2nd and 3rd eigenvectors representing the closing motion of the α-helical subdomain on the β-sheet subdomain. The first three eigenvectors of both VPL57 and VPL58 describe almost 40% of the collective motion of the complex.

Interestingly, the crystal structures of EBOV VP35-RNA complex suggest that the Phe328-Leu336 loop can be involved in protein-protein lateral contacts between EBOV VP35 units during the recognition process of double stranded viral RNA. This process have been reported to be critical for inhibition of interferon induced antiviral activity and EBOV virulence,[34] thus affecting its dynamics might perturb also this step of the suppression of host immunity by EBOV. On the other hand, Phe328-Leu336 loop seems not to be of crucial importance for RNA binding in other filoviruses, as discussed previously.[35]

Conclusions

In this work we studied EBOV VP35 and its inhibitors from a computational point of view. We performed MD simulation of apo-protein and complexes starting from the crystallographic structures and we performed CAS, obtaining a very good agreement with

the structural and biochemical experimental results. We deepened the study of this systems from the point of view of the intermolecular interactions, identifying new hot spots in addition to those experimentally determined. Then we performed the RMSF bootstrap analysis, in order to identify the most mobile residues and to highlight the difference between the apo-protein and the complexes, in particular we identified three protein segments whose behavior was different in the apo-protein compared to the complexes.

To characterize the RMSF analysis, in order to analyse the global motion of the systems, we performed ED analysis, obtaining the collective and most important motion for every system. We noticed that ligands hinder the typical closing motion of the β sheet-subdomain towards the α helical-subdomain of the apo-form. The binding of some ligands, like VPL57 VPL58 and VPL60 exalt the movement of the loop Phe328-Leu336.

Notes and references

1. H. Feldmann and T. W. Geisbert, *The Lancet*, 2011, **377**, 849-862.
2. C. M. Bosio, M.J. Aman, C. Grogan, R. Hogan, G. Ruthel, D. Negley et al., *J. Infect Dis*, 2003, **188**, 1630-1638.
3. M. Bray, T. W. Geisbert, *Int J Biochem Cell Biol*, 2005, **37**, 1560-1566.
4. S. P. Reid, L. W. Leung, A. L. Hartman, O. Martinez, M. L. Shaw, C. Carbonelle, et al., *J Virol*, 2006, **80**, 5156-5167.
5. S. P. Reid, C. Valmas, O. Martinez, F. M. Sanchez, C. F. Basler, *J Virol*, 2007, **81**, 13469-13477.
6. Z. Feng, M. Cerveny, Z. Yan and B. He, *J Virol*, 2007, **81**, 182-192.
7. M. Schümann, T. Gantke and E. Mühlberger, *J Virol*, 2009, **83**, 8993-8997.
8. J. Haasnoot, W de Vries, E-J. Geutjes, M. Prins, P. de Haan et al., *PLoS Pathog*, 2007, DOI: 10.1371/journal.ppat.0030086
9. W. B. Cárdenas, Y-M. Loo, M. Gale Jr., A. L. Hartman, C. R. Kimberlin, L. Martinez-Sobrido, E. O. Saphire and C. F. Basler, *J Virol*, 2006, **80**, 5168-5178.
10. C. S. Brown *et al.*, *J. Mol. Biol.*, 2014, **426**, 2045-2058.
11. D. W. Leung, N. D. Ginder, D. B. Fulton, J. Nix, C. F. Basler, R. B. Honzatko and G. K. Amarasinghe, *PNAS*, 2009, **106**, 411-416.
12. Bale *et al.*, PLoS Pathog., 2012, **8**, DOI: 10.1371/journal.ppat.1002916.
13. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, J. *Chem. Theory Comput.*, 2008, **4** , 435–447.
14. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950-1958.
15. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of Computational Chemistry*, 2004, **25**, 1157-1174.
16. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.,* **79**, 926-935.
17. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaje, *J. Comput. Chem., 1997,* **18**, 1463-1472.
18. G. Bussi, D. Donadio and M. Parrinello*, J. Chem. Phys.*, 2007, **126**, DOI: http://dx.doi.org/10.1063/1.2408420
19. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
20. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.,* 1993, **98**, 10089-10092.
21. S. Hubbard and J. Thornton, Naccess, The University of Manchester, UK, 1992-6
22. P. A. Kollman *et al., Acc. Chem. Res.*, 2000, **33**, 889–897.
23. S. Huo, I. Massova and P. A. Kollman, *J. Comput. Chem.*, 2002, **93**, 15-27;
24. S. Pieraccini, R. De Gonda and M. Sironi, *Chemical Physics Letters*, 2011, **517**, 217-222.
25. S. Huo *et al.*, J. Comput. Chem. 2002, **23**, 15-27.
26. I.S. Moreira *et al.*, J. Comput. Chem. 2007, **28**, 644-654.
27. N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *PNAS*, 2001, **98**, 10037-10041.
28. I. Massova and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 8133–8143.
29. A. Mitra and D. Sept, *Biophysical Journal*, 2008, **95**, 3252-3258.
30. A. Amadei, A. B. M. Linssen and H. J. C. Berendsen, *Proteins*, 1993, **17**, 412-425
31. W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33-38.
32. T. Huo, J. Wang, Y. Li and W. Wang, J. Chem. Inf. Model., 2011, **51**,69-82.
33. I.S. Moreira *et al.* Theor. Chem Acc., 2007, **117**, 99-113.
34. D.W. Leung *et al.*, Nat Struct Mol Biol., 2010, **17**, 165-172.
35. Ramanan *et al.*, PNAS, 2012, **109**, 20661-206666.

Figures and table

Residue Number	ΔΔG / kcal mol^{-1}								
	GA017	GA246	VPL27	VPL29	VPL42	VPL48	VPL57	VPL58	VPL60
Lys 222	**9.80**	**4.61**	**4.53**	**3.35**	**4.41**	**5.14**	**3.39**	**2.98**	**3.27**
Arg 225	**9.19**	**17.28**	**6.22**	**4.80**	**7.47**	**8.33**	**5.59**	**6.27**	**6.24**
Gln 241	0.24	1.10	0.65	0.28	-0.91	0.72	0.80	-0.23	-0.26
Gln 244	**2.75**	**3.26**	**4.56**	**1.96**	**4.29**	**2.77**	**4.54**	**4.14**	**4.07**
Val 245	0.51	0.61	0.77	0.72	0.74	0.65	0.63	0.62	0.69
Cys 247	0.21	0.35	0.12	0.32	-0.01	0.14	0.21	0.05	0.21
Lys 248	**9.91**	**8.73**	**12.02**	**14.83**	**11.16**	**9.42**	**14.39**	**12.63**	**14.54**
Leu 249	0.24	0.29	0.60	0.56	0.51	0.67	0.64	0.18	0.62
Lys 251	**12.96**	**11.22**	**8.00**	**11.14**	**8.92**	**10.61**	**8.43**	**10.04**	**10.19**
Asp 252	-1.28	-0.85	-2.48	-4.73	-3.51	-2.23	-5.53	-1.97	-4.28
Asp 289	1.09	1.26	1.05	1.04	1.11	1.05	0.99	1.04	0.98
Val 294	0.01	0.03	0.04	0.04	0.03	0.04	0.04	0.06	0.03
Ile 295	**3.93**	**5.29**	**4.78**	**4.08**	**4.49**	**5.24**	**4.97**	**4.70**	**4.76**
His 296	-0.13	-0.19	-0.18	-0.10	-0.10	0.04	-0.05	-0.18	-0.16
Ile 297	0.23	0.96	0.80	0.78	0.73	1.13	0.87	1.10	0.87
Asp 302	-0.83	-0.97	-1.07	-0.81	-1.03	-0.41	-1.05	-0.93	-1.27
Ile 303	0.54	0.22	0.06	0.25	0.22	0.29	0.38	0.32	0.15
Phe 328	0.38	1.09	0.77	0.83	0.85	1.23	0.67	0.69	0.84

Table 1 CAS values for protein-ligand complexes. The hot spots are in bold.

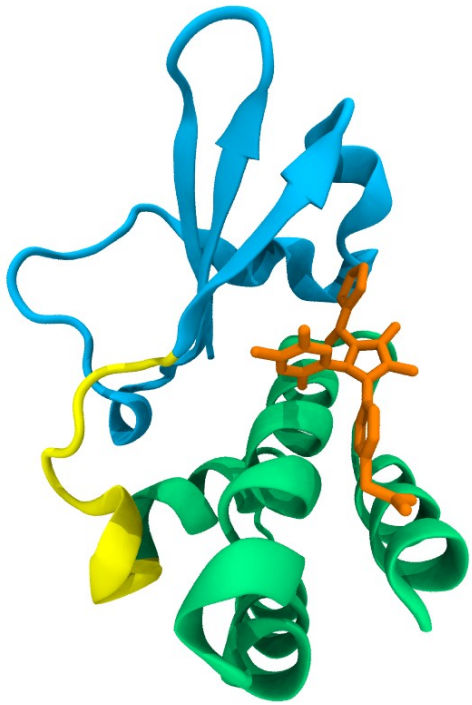


Figure 1 VP35 in complex with one of the ligands (GA017). The α-helical subdomain (residues from Ala221 to Arg283) is represented in green and the β-sheet subdomain (residues from Val294 to Ile340) is represented in blue. The binding mode is the same for all ligands.

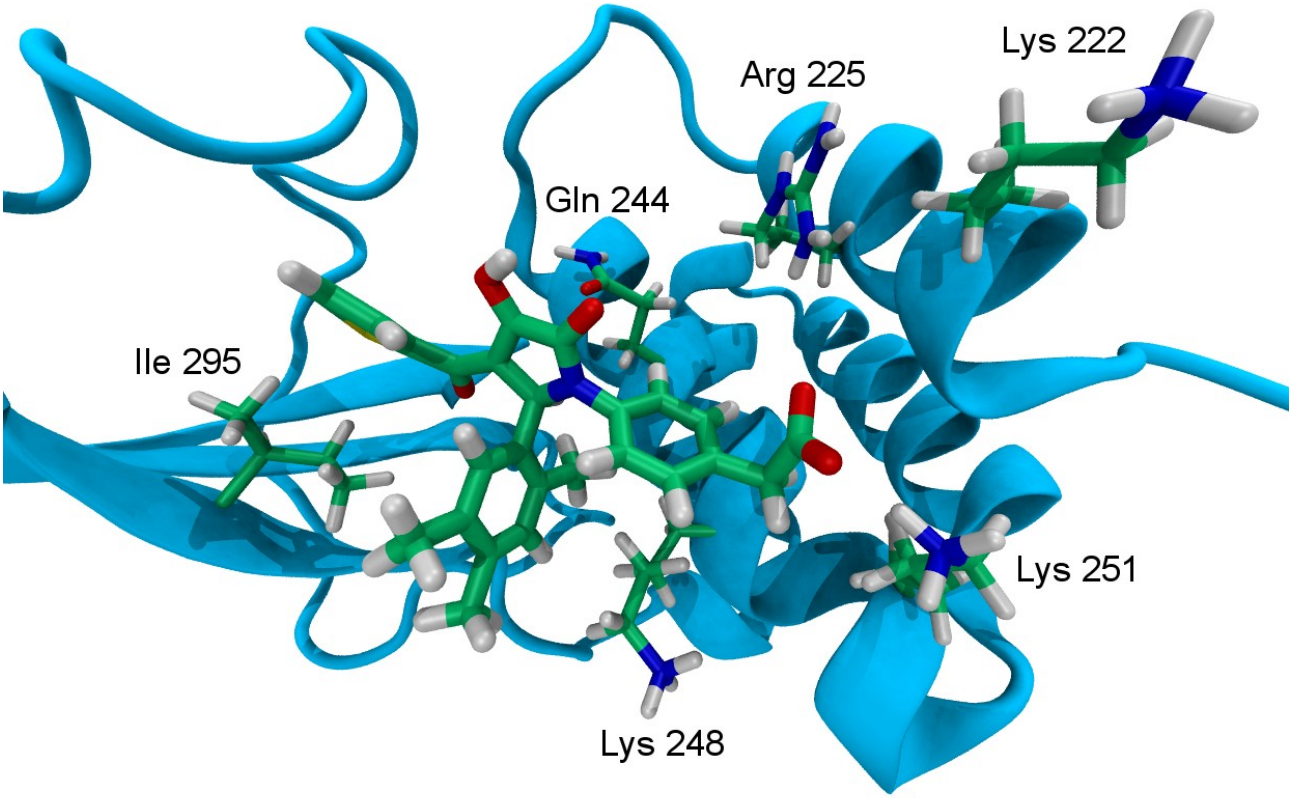


Figure 2 An overview of the hot-spot residues in GA017-VP35 complex. These residues are common to all the complexes studied.

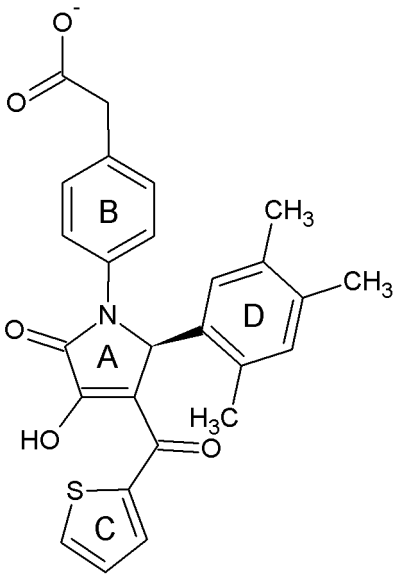


Figure 3 The structure of GA017. The pyrrolidinone scaffold is common to all the molecules. Molecular structures of all the simulated ligands can be found in Supplementary Informations (Fig. S1).

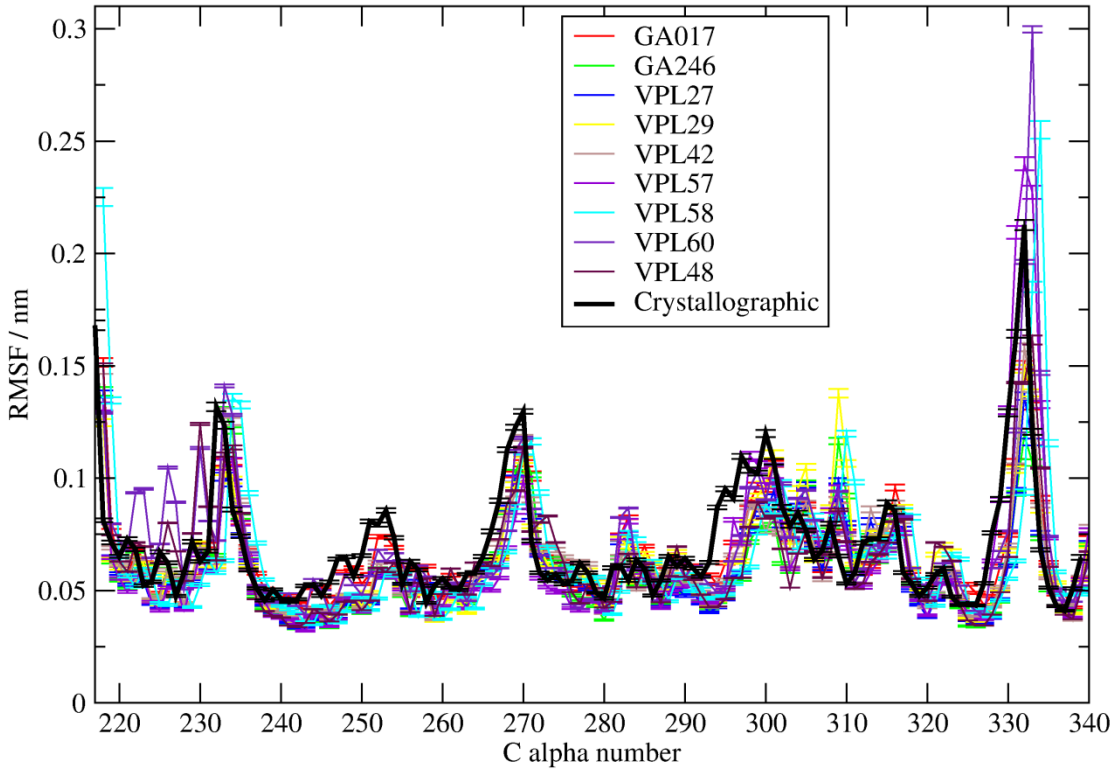


Figure 4 RMSF bootstrap analysis. Three protein segments have been further analyzed, due to their larger flexibility: Ile246-Ser253, Pro293-Ile297 and Phe328-Leu336 (see text). A larger version of this figure can be found in Supplementary Informations (Fig. S11).

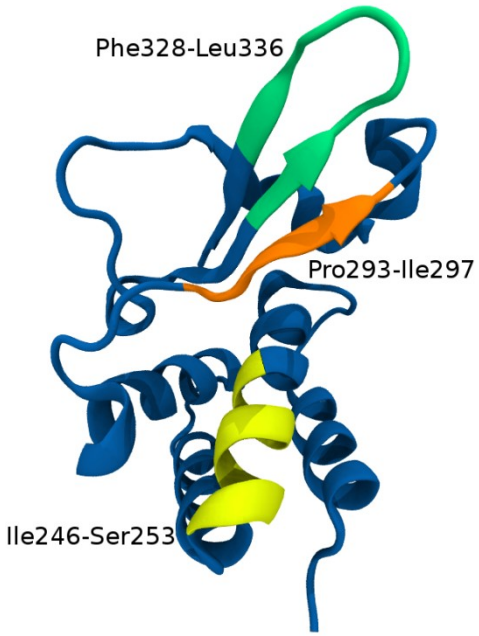


Figure 5 Flexible domains of VP35.

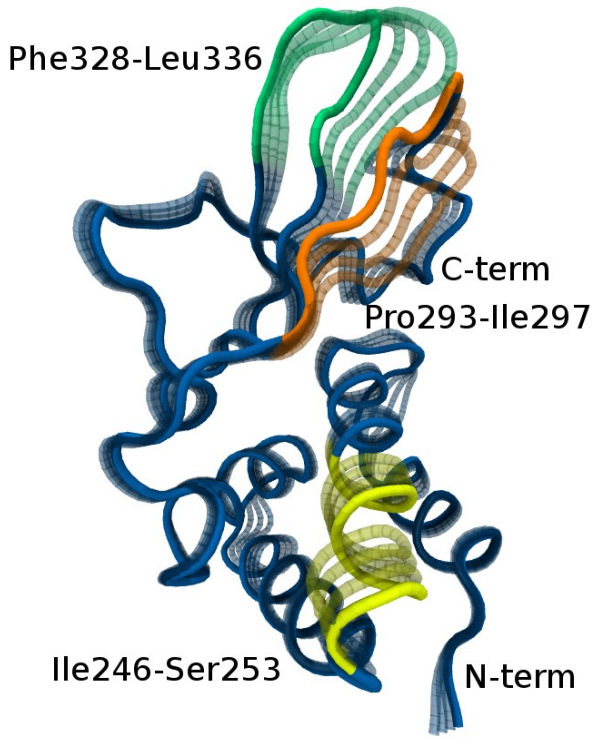


Figure 6 Schematic representation of the projection of the MD simulation on the 1st eigenvector extracted from the ED of the apo-protein. Highlighted in orange and green we can see the closing motion of the β-sheet subdomain towards residues Ile 246-Ser 253 belonging to the α-helical subdomain, highlighted in yellow.

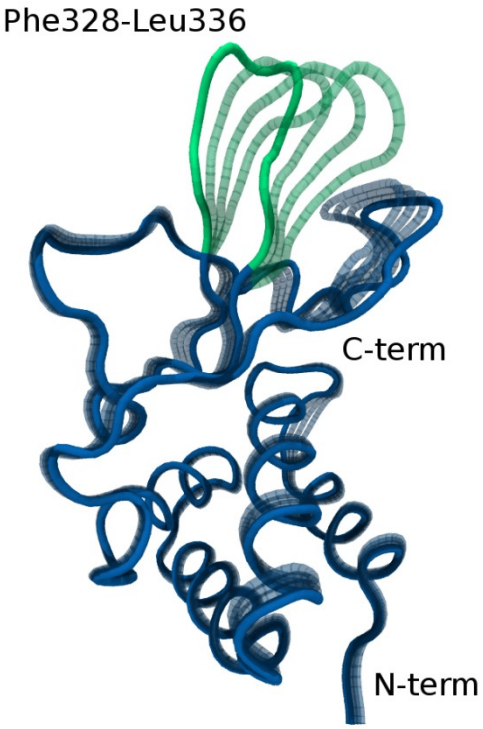


Figure 7 The projection of the MD simulation on the 1st eigenvector of VPL57-VP35 complex. Highlighted in green we can see the intense motion of the Phe328-Leu336 loop. The characterizing closing motion of the apo-protein is not present in this case.