Computer aided design and NMR characterization of an oligopeptide targeting Ebola virus VP24 protein

Federico Dapiaggi[a], Stefano Pieraccini[a,b], Donatella Potenza[a], Francesca Vasile[a], Helena Macut[c], Sara Pellegrino[c], Alessandro Aliverti[d], Maurizio 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 (INSTM), CNR, and INSTM, UdR Milano, Via Golgi 19, 20133, Milano, Italy

c) DISFARM-Dipartimento di Scienze Farmaceutiche, Sezione Chimica Generale e Organica "A. Marchesini", Via Venezian 21, 20133, Milano, Italy

d) Dipartimento di Bioscienze, Via Celoria 26, 20133, Milano, Italy

**To whom correspondence should be addressed: Stefano.pieraccini@unimi.it , Maurizio.sironi@unimi.it*

Abstract

The Ebola virus Viral Protein 24 (VP24) inhibits interferon signaling through its interaction with the human protein Karyopherin, thus impairing the immune response of the host against the infection and increasing its rate of diffusion into the organism and its lethality. This makes of VP24 a potential pharmacological target, as the inhibition of its interaction with Karyopherin could reduce Ebola virus virulence. In this work, we carried out an atomic level study of the network of interactions between VP24 and Karyopherin using molecular dynamics and computational alanine scanning. Modeling the VP24-Karyopherin complex allowed us to identify the amino acid residues responsible for protein-protein binding and led to the identification of a nonapeptide with VP24 binding potential. Subsequently, the ability of this peptide to actually bind VP24 in solution has been assayed by Saturation Transfer Difference NMR and Circular Dichroism (CD). Experimental and molecular modeling data concerning the VP24-peptide complex have been compared and putative peptide binding site and mode are discussed.

Introduction

Ebola virus (EBOV) epidemics have repeatedly burst in equatorial Africa in the last decades, highlighting the need of new therapies against this disease. EBOV virus belongs to the *Filoviridae* family of single stranded, non-segmented negative-sense RNA viruses. EBOV is the etiologic agent of a haemorrhagic fever with a very high human fatality rate, ranging from 50% to near 90%.[1] The virulence and high lethality of this virus are due to diverse factors, in particular to its ability to inhibit both the innate immune response in the early stages of infection and the subsequent adaptive specific immune responses of the host organism.[2,3] This is done 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, so there is a very limited number of viral targets available. Moreover, 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 themselves or with human proteins. This makes protein-protein interactions targeting a necessary strategy for contrasting EBOV infection. Recently, different research groups focused their attention on Viral Protein 35, a multitasking and attractive target in the Ebola proteome[6-8]. In this work, we consider Viral Protein 24 (VP24), which acts as an inhibitor of interferon (IFN) signaling, as a potential target. Immune response of the organism to a viral infection is due to interferons, through the phosphorylation of the STAT (signal transducer and activator of transcription) protein operated by the Janus kinase[9]. The phosphorylated STAT is recognized by a subset of Karyopherin α (KPNA) family of nuclear transport factor. This STAT-KPNA complex is transported in the cell nucleus where it induces the expression of ISG gene, that has an antiviral activity.[10,11] VP24 acts in a cell intrinsic manner inhibiting IFN signaling and making cells refractory to IFN as it forms a complex with KPNA, thus making it unavailable for the STAT transportation into the cell nucleus.[12] Inhibition of VP24-Karyopherin interaction is a possible therapeutic strategy to reduce Ebola virulence. Targeting protein-protein interactions (PPIs) is a difficult task, due to the inherent features of protein-protein contact surfaces, which are often flat and lack well defined pockets, and to their variability[13,14]. Nevertheless, in recent years many encouraging results have been obtained, showing that exploiting PPIs as therapeutic targets is a viable, albeit challenging option[15,16], and several PPIs targeting molecules, both of peptidic[17-19] and non peptidic nature[20-22], have been described. Among the molecular modeling techniques aimed to study PPIs, computational alanine scanning[23] (CAS) is a widely used one. It consists in the evaluation of the difference in the binding free energy of the proteins forming a complex upon mutation of each of the interfacial residues into alanine. This evaluation is performed with *a posteriori* free energy calculation approaches, such as the MM/PBSA[24], on snapshots extracted from a molecular dynamics trajectory. This kind of approach has been successfully used to identify peptides acting as PPIs inhibitors and corresponding to protein subsets located at the protein-protein

interface[25-27], as well as to describe the interactions between protein surface and small molecule binders[28,29]. Among the experimental techniques suitable for studying transient protein-peptide complexes the Nuclear Magnetic Resonance spectroscopy (NMR) is one of the most powerful and versatile method, well suited to provide a detailed description of the receptor-ligand interactions[30,31]. In particular Saturation Transfer Difference (STD-NMR) technique is one of the most widespread NMR methods to study the interactions between small ligands and macromolecular receptors[32]. In this work we performed an atomic level mapping of the interactions between VP24 and Karyopherin, employing molecular dynamics simulations in explicit solvent and free energy calculations, and identified a small subset of residues at the VP24-KPNA interface responsible for protein complex formation. In particular, such critical residues belonging to Karyopherin were used as the basis for the design of a nine residues long peptide potentially able to interact with VP24, thus competitively inhibiting its interaction with Karyopherin. To evaluate whether the identified peptide could retain its ability to interact with VP24 even when extracted from its protein environment, molecular dynamics simulation of the peptide-VP24 complex were performed, as well as CAS, in order to evaluate the structural and energetic behavior of the peptide. The interaction between VP24 and the identified peptide was then experimentally investigated through circular dichroism (CD) and advanced NMR techniques. The interactions between the Ebola receptor and the nonapeptide ligand are governed by weak forces including Van der Waals interactions, hydrogen bonding and hydrophobic associations. In this case, the ligand-based NMR approach represent an optimal angle of observation, because it permits the analysis of interactions governed by weak equilibrium dissociation constants (in the micromolar to the millimolar range), focusing on the ligand[33]. We confirmed the binding of the peptide with VP24 protein and mapped the interaction epitope by determining the ligand regions in contact with the receptor.

Materials and Methods

Molecular dynamics

The crystallographic structure of the complex VP24-KPNA was obtained from the Protein Data Bank (PDB code: 4U2X[12]), and the residue numbering used here is referred to such structure. Molecular dynamics simulations were carried out with the GROMACS 4.5.3[34] package using explicit solvent and periodic boundary conditions. The AMBER99SB-ildn[35] force field was used. Every system was solvated with TIP4P[36] waters and neutralized with Cl^- or Na^+ ions to reach neutrality. The LINCS algorithm[37] was employed to constraint all bonds to their equilibrium length, allowing a time step of 2 fs. The system was submitted to 10'000 steps of geometry optimization with the steepest descent method. Afterwards it was equilibrated for 200 ps in NVT conditions (T=300 K) and subsequently for 200 ps in NPT conditions, in order to equilibrate systems density. For the VP24-KPNA complex a 50 ns molecular dynamics was performed in NPT conditions (1 bar, T=300 K).

For each of the peptide-VP24 complexes, a 100 ns molecular dynamics was run in the aforementioned conditions. Temperature and pressure were kept constant to their reference value using the velocity rescale algorithm[38] and the Berendsen barostat[39], respectively. A 14 Å cutoff was applied for non-bonded interactions and the Particles Mesh Ewald algorithm[40] was employed to calculate long range electrostatic interactions.

Computational alanine scanning

500 snapshots were extracted from the last 20 ns of the dynamics of the VP24-K complex (one snapshot every 40 ps). Solvent accessible surface area (SASA) of each residue was calculated using Naccess[41]. The ΔG of binding was calculated with the MM/PBSA approach as implemented in the GMXPBSA 2.0 suite[42-45]. This protocol implicitly assumes that point mutations in the protein do not significantly affect its conformation. The validity of this assumption in computational alanine scanning has been widely confirmed in the literature, when applied to protein-protein interactions[46-47]. A dielectric constant of 2 was chosen for protein interior. The same protocol was employed for the computational alanine scanning for each VP24-peptide complex.

VP24 purification

The portion of Ebola virus VP24 protein comprising residues 1-233 and lacking the structurally disordered 18-residue C-terminal region was produced in *Escherichia coli* strain Rosetta(DE3) fused to an N-terminal poly-His sequence to facilitate its purification, using the plasmid construct based on pET-46 EkLIC vector described elsewhere[48,49]. Bacterial cultures grown under agitation in 2×YT medium, supplemented with 30 mg/L chloramphenicol and 100 mg/L ampicillin, were induced with 0.2 mM isopropil-β-D-1-tiogalattopiranoside for 16 h at 25 °C. Cells (typically 15 g wet weight) were suspended in 30 mL of 50 mM Na-phosphate, pH 7.4, containing 200 mMNaCl, 5 mM imidazole (Solution A), added with Complete® EDTA-free protease inhibitor cocktail (Sigma-Aldrich), and disrupted by sonication. After removal of cell debris by centrifugation, the crude extract was loaded on two 5 mL HisTrap HP cartridges (GE Healthcare) connected in series, equilibrated with Solution A. After extensive washing, initially with Solution A and then bringing imidazole concentration to 25 mM, VP24 elution was obtained by a 25-500 mM imidazole concentration gradient over 80 mL in Solution A. After protein concentration by ultrafiltration to about 25 mg/mL, solvent was changed to 50 mM Na-phosphate, 500 mM NaCl by gel filtration on a HiTrap desalting column (GE Healthcare). Purified VP24 (typically 80 mg) was subdivided into aliquots and stored at -20 °C. The extinction coefficient used to quantitate recombinant VP24 was determined by measuring its absorbance at 280 nm in 50 mM Na-phosphate, pH 7.4, 200 mM NaCl, in the absence and in the presence of 5 M guanidine hydrochloride. Then, from the ε_{280} value computed through the ExPASy server for the protein under denaturing conditions[50] a ε_{280} of 31,390 $M^{-1}cm^{-1}$ for native VP24 1-233 fragment was obtained.

Peptide synthesis

The peptide RS (AYGLDKIEF) was prepared by microwave-assisted solid phase synthesis[51] based on Fmoc chemistry on Fmoc-Rinkamide resin (0.57 meq/g substitution), using a fivefold molar excess of 0.2 M Fmoc-protected amino acids dissolved in N-methyl pyrrolidinone, and using HOBT/HBTU/DIEA (5 : 5 : 10 eq) as activators. Coupling reactions were performed for 5 min at 40 W with a maximum temperature of 75 °C. Deprotection was performed in two stages using 20% piperidine in dimethylformamide (5 and 10 min each). The peptide was acetylated using Ac_2O (10 eq) and DIEA (10 eq) in DCM/DMF (1:3). The coupling reaction was performed twice for 0.5 h each time. Cleavage from the resin was performed using 10 mL of Reagent K (trifluoroacetic acid/phenol/water/thioanisole/ 1,2-ethanedithiol; 82.5 : 5 : 5 : 5 : 2.5) for 180 min. Following cleavage, the peptide was precipitated and washed using ice-cold anhydrous ethyl ether. The peptide was purified by RP-HPLC using a gradient elution of 5–70% solvent B (solvent A: water/acetonitrile/trifluoroacetic acid 95 : 5 : 0.1; solvent B: water/acetonitrile/trifluoroacetic acid 5 : 95 : 0.1) over 20 min at a flow rate of 20 mL/min. The purified peptide was freeze-dried and stored at 0 °C.

Circular Dichroism

CD spectra were collected using a Jasco J-810 spectropolarimeter. Peptide RS stock solution was prepared in 50 mM PBS (100 µM, pH 7.4). VP24 stock solution was prepared in 50 mM PBS and 200 mM NaCl (95 µM, pH 7.4). Spectra were obtained from 200 to 260 nm with a 0.1 nm step and 1 s collection time per step, taking three averages. The spectrum of the buffer was subtracted to eliminate interference from cell, solvent, and optical equipment. The CD spectra were plotted as mean residue ellipticity (degree x cm2 x dmol-1) versus wave length (nm). Noise-reduction was obtained using a Fourier-transform filter program from Jasco.

Nuclear Magnetic Resonance

The peptide was characterized by one- and two-dimensional NMR experiments. NMR spectra were collected on samples containing 2 mM peptide in 20 nM phosphate buffer (containing 10% D2O). The assignment[52] are reported in the Table S2 of Supplementary Information. The spectra were acquired with a Bruker Avance 500 MHz instrument at 298 K. TOCSY spectra, with a mixing time of 60 ms, allow reaching a complete assignment of all the spin systems of the amino acids. NOESY experiments with mixing times of 700 ms and ROESY with mixing time of 200 ms were used to evaluate the spatial correlations. In 1D, TOCSY, NOESY and ROESY experiments, solvent suppression was achieved using the excitation sculpting pulse sequence. In STD experiments 5000 Hz (10 ppm) and -100 Hz (-0.2 ppm) were used as on-resonance irradiation, while 100000 Hz (200 ppm) was chosen as off-resonance. A train of Gauss shaped pulses of 49 ms length each achieved a selective pre-saturation of the protein. The experiments were performed with 2.94 s of total saturation. In the STD experiment water suppression was achieved by use of the WATERGATE 3–9–19 pulse sequence. STD experiment was acquired also in absence of protein in order to exclude artifacts. In the analysis of the binding,

the final concentrations of peptide and protein were about 1.1 mM and 22 μM, respectively, resulting in a 50:1 ligand-target ratio.

Results and Discussion

Computational alanine scanning

A molecular dynamics of 50 ns was performed on the VP24-KPNA complex. Computational alanine scanning (CAS) was carried out in the last 20 ns of the dynamics when the system was fully equilibrated as shown by RMSD plot (Fig. S1-S2). In general, a few nanoseconds long trajectories allow a sufficient sampling for an accurate binding free energy calculation with the MM/PBSA approach and consequently for CAS purposes[53]. The protein-protein interface is defined as the ensemble of amino acids whose solvent exposed surface area has a nonzero variation upon complex formation and it is composed of 56 residues. Residues whose mutation led to a variation in binding free energy (ΔΔG) greater than 2 kcal mol^{-1} are defined as hot-spots. CAS results are summarized in Table 1. Hot spots are evenly distributed between VP24 and KPNA, which contain 10 and 11 hot spots respectively (Fig.1a).

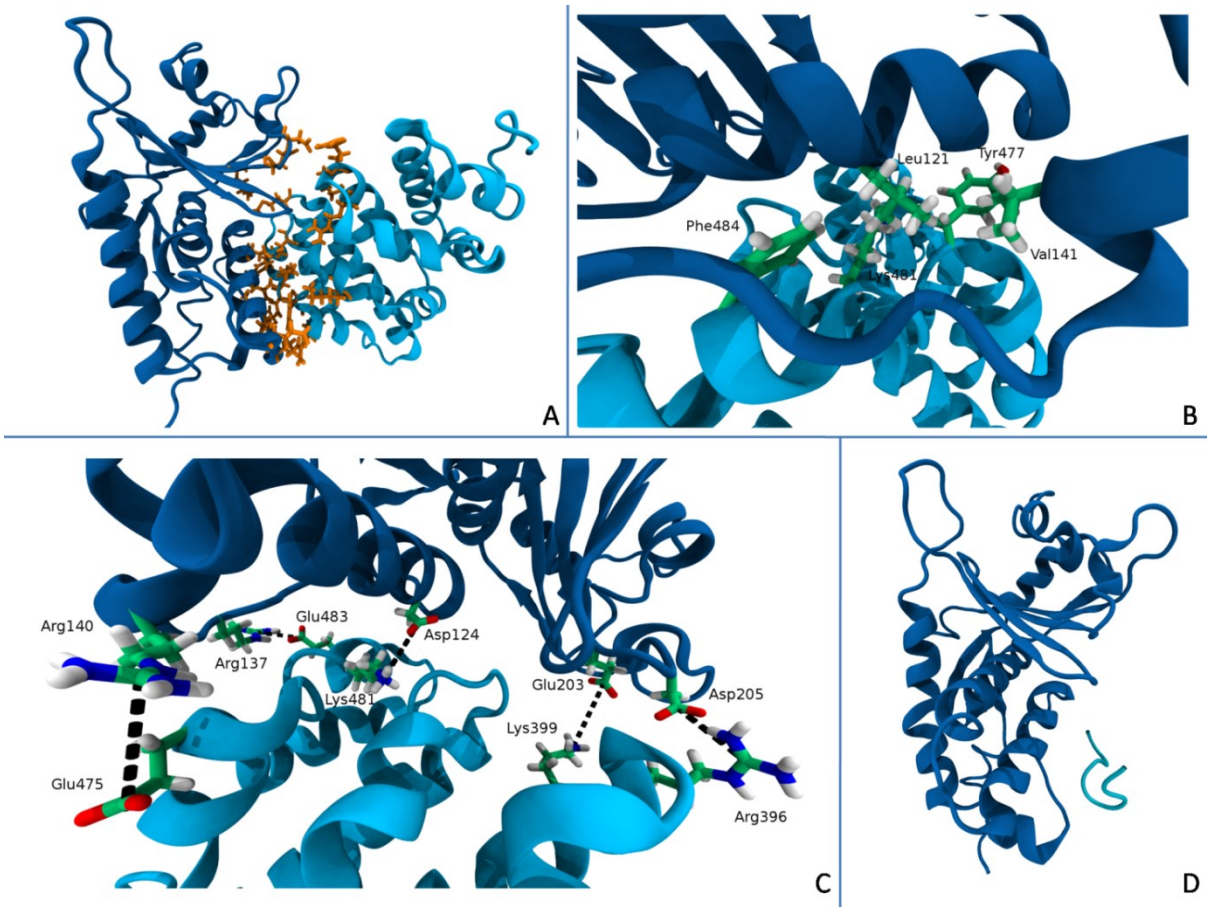

Fig. 1 A) VP24 (blue)-KPNA (cyan) complex. The hot-spots arising from the CAS analysis are highlighted in orange. B) The hydrophobic cluster at the VP24-KPNA interface. C) Overview of the five salt bridges at the

VP24-K interface. D) VP24 (blue) in complex with the RS peptide (cyan) extracted from KPNA. Figures were rendered with VMD[58].

The two proteins exhibit a high degree of electrostatic complementarity, as can be inferred from electrostatic potential maps (Fig. S3-S4). A more detailed insight into the network of interactions at the interface that contribute to VP24-KPNA complex formation can be achieved by the analysis of the local environment of the hot spots during the simulation. As we can see from Table 1, since most of the hot-spots are either charged or polar residues, most of the interactions within the VP24-KPNA complex are salt bridges and hydrogen bonds, although hydrophobic hot spots and non-polar interactions are also observed. In particular, Arg140 and Glu475' (residues belonging to KPNA are labeled with a prime on their sequence number) form a salt bridge that is populated along 88% of the MD trajectory, while salt bridges between Glu203 and Lys 399', between Asp205 and Arg396' and between Asp124 and Lys481' are populated during 85%, 78% and 64% of the trajectory respectively (Fig.1c). Residue Asp124 is also involved in a hydrogen bond between its carboxyl group and Thr434' hydroxyl, which is observed during 75% of the simulation, while Arg137 and Asp480' backbones form a hydrogen bond for 74% of the trajectory. A small hydrophobic cluster, comprising Leu121, Val141 and Tyr477', Phe484' and the aliphatic portion of the sidechain of Lys481' has been also observed and is steadily formed during the whole simulation (Fig. 1b). An interaction table for the complex VP24-KPNA can be found in supplementary material (Table S3). Table 1 also shows the presence of some residues exhibiting a largely negative ΔΔG (i.e. residues whose mutation into alanine increase the binding energy between the proteins). These are Glu88, Glu394', Asp431' and Asp437'. Inspecting the electrostatic potential on VP24 and KPNA surface, it may be observed (Fig. S5-S6) that all these negatively charged residues, when the complex is formed, actually face a negative potential area on the binding partner. This suggest a somehow suboptimal affinity of VP24 for KPNA, which deserves further attention and should be investigated by molecular biologists. Stemming from these geometric and energetic features of the VP24-KPNA interface, we identified two KPNA segments comprising a set of hot spot residues that are close to one another not only in the protein tridimensional structure, but also in its sequence. The binding capability of an isolated peptide is not guaranteed in principle, even if it contains several hot spots, because it may undergo major structural rearrangements when isolated from the parent protein and because of the complexity of the interaction networks that lead to protein complex formation. The first sequence we considered, comprising the hot spot residues Arg396', Arg398' and Lys399', ranges from Ala393' to Glu400'. A 100ns MD simulation of this peptide in complex with VP24 shows large conformational changes (Fig. S12), leading to a partial detachment of the peptide from the protein. In particular, the salt bridge interactions involving the former hot-spots are completely lost. For these reasons the peptide seems unable to interact with VP24, thus it was no longer considered. The second sequence comprises residues ranging from Glu474' to Phe484'. A 100ns MD simulation and a CAS of the peptide-VP24 complex were performed in order to evaluate the binding capability of the selected sequence. Both Glu474' (hot spot in VP24-KPNA complex) and Glu475' lost their

hot spot character showing negative ΔΔG values (Fig. S11), while all the remaining hot spots are conserved. Based on these results we decided to shorten the sequence discarding Glu474' and Glu475', obtaining a less charged peptide maintaining a very high density of hot spots, namely Tyr477', Asp480', Lys481' and Phe484'. We inferred that a peptide sequence corresponding to this KPNA subsequence could bind to VP24, thus interfering with VP24-KPNA complex formation. The selected peptide (from now on named RS, Fig. 1d) involves residues ranging from Ala476' to Phe484' and has the sequence AYGLDKIEF. To investigate whether the identified peptide could retain its ability to bind to VP24, even when extracted from its protein environment, we ran a 100 ns long control molecular dynamics simulations of the VP24–RS complex. The protein-peptide complex showed a good structural stability during the whole simulation, with protein binding site and peptide binding mode well conserved. On the other hand, its helical secondary structure was progressively relaxed and eventually lost during the simulation (Fig. S7).

Mutation	$\Delta\Delta G$ / $kcal\ mol^{-1}$	err / $kcal\ mol^{-1}$
GLU88	-10.99	0.79
GLU113	**6.16**	**0.79**
LEU121	**6.28**	**0.79**
ASP124	**4.35**	**0.79**
TRP125	0.96	0.79
LEU127	0.14	0.79
THR128	-0.64	0.79
THR129	1.53	0.79
ASN130	0.24	0.79
THR131	-0.21	0.79
PHE134	1.84	0.79
ASN135	**3.30**	**0.79**
MET136	1.62	0.79
ARG137	**23.84**	**0.81**
THR138	-2.89	0.79
GLN139	0.29	0.79
ARG140	**9.94**	**0.79**
VAL141	**2.65**	**0.79**
GLN184	-1.55	0.79
ASN185	1.15	0.74
HIS186	0.14	0.79
LEU201	1.27	0.79
GLN202	-0.07	0.79
GLU203	**10.37**	**0.72**
ASP205	**16.31**	**0.76**
SER207	0.24	0.79
ASN210	-0.38	0.79
LYS218	**11.13**	**0.79**

Mutation	$\Delta\Delta G$ / $kcal\ mol^{-1}$	err / $kcal\ mol^{-1}$
LEU390	0.10	0.79
GLN391	0.41	0.79
GLU394	-13.16	0.79
PHE395	1.43	0.76
ARG396	**21.04**	**0.79**
ARG398	**18.80**	**0.74**
LYS399	**23.07**	**0.84**
LYS427	**16.67**	**0.79**
ASP431	-14.81	0.79
LEU433	0.50	0.79
THR434	**6.09**	**0.79**
VAL435	1.19	0.79
MET436	**3.63**	**0.79**
ASP437	-13.57	0.74
GLU474	**7.91**	**0.79**
GLU475	1.98	0.79
TYR477	**6.31**	**0.79**
LEU479	0.62	0.79
ASP480	**17.56**	**0.79**
LYS481	**28.90**	**0.81**
GLU483	0.79	0.79
PHE484	**6.02**	**0.79**
LEU485	0.55	0.79
SER487	-1.53	0.79
HIS488	1.15	0.79
GLU489	-4.44	0.76
ILE501	-0.45	0.79
PHE505	1.00	0.79

Tab.1 Computational Alanine Scanning data with standard errors. Left column refers to VP24 residues, right column to KPNA residues. Hot spots are reported in bold.

Subsequently, a CAS was carried out in order to verify if the previously identified KPNA hot spots were conserved also in the VP24-RS complex. As shown in Fig. 2 and Table S1, not only all the hot spots are conserved, but also a new one, i.e. Glu483', can be identified. This residue forms a salt bridge with Arg137 which is observed in 90% of the trajectory of the VP24-RS complex. Interestingly, this salt bridge, albeit present also in the VP24-KPNA complex, was observed only in 31% of the trajectory.

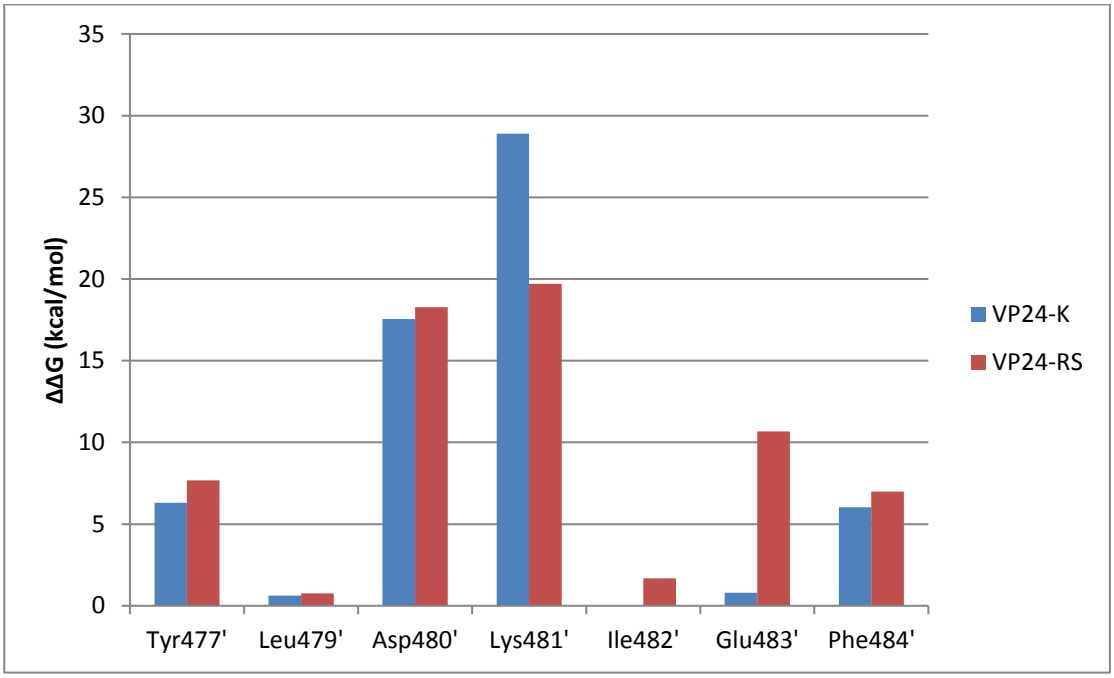

Fig. 2 Computational Alanine Scanning results for residues ranging from Tyr 477' to Phe 484' belonging to Karyopherin (blue bars) or to the RS peptide (red bars).

Nuclear Magnetic Resonance

In order to experimentally verify if RS can actually bind VP24 in solution, NMR experiments were performed. First, the capped-peptide was analyzed in terms of conformational features by NOESY (500ms) and ROESY (200ms), observing no medium and long-range correlations. Only sequential correlations could be found, indicating that the peptide exists in random conformation. tr-NOESY and tr-ROESY experiments showed no definite secondary structure in the peptide even upon addition of VP24. These observations mirror the behavior of RS in MD simulations, both alone and in complex with VP24 (Fig. S8-S9). In this paper, STD technique was used as an epitope mapping device to describe the VP24-peptide interactions. The method is based on the transfer of saturation from the protein to the bound peptide which in turn, by exchange, is moved into solution where it is detected. During the period of saturation, the magnetization gradually moves from the protein to the protons of the peptide when the ligand binds to the target. For transient interactions with rapid exchange[54,55], the ligand polarization in the bound state is transferred to the free state where the

saturation accumulates during the irradiation time of the experiment. The STD spectrum shows only the signals of peptide that were in close contact with the protein (Fig. 3).

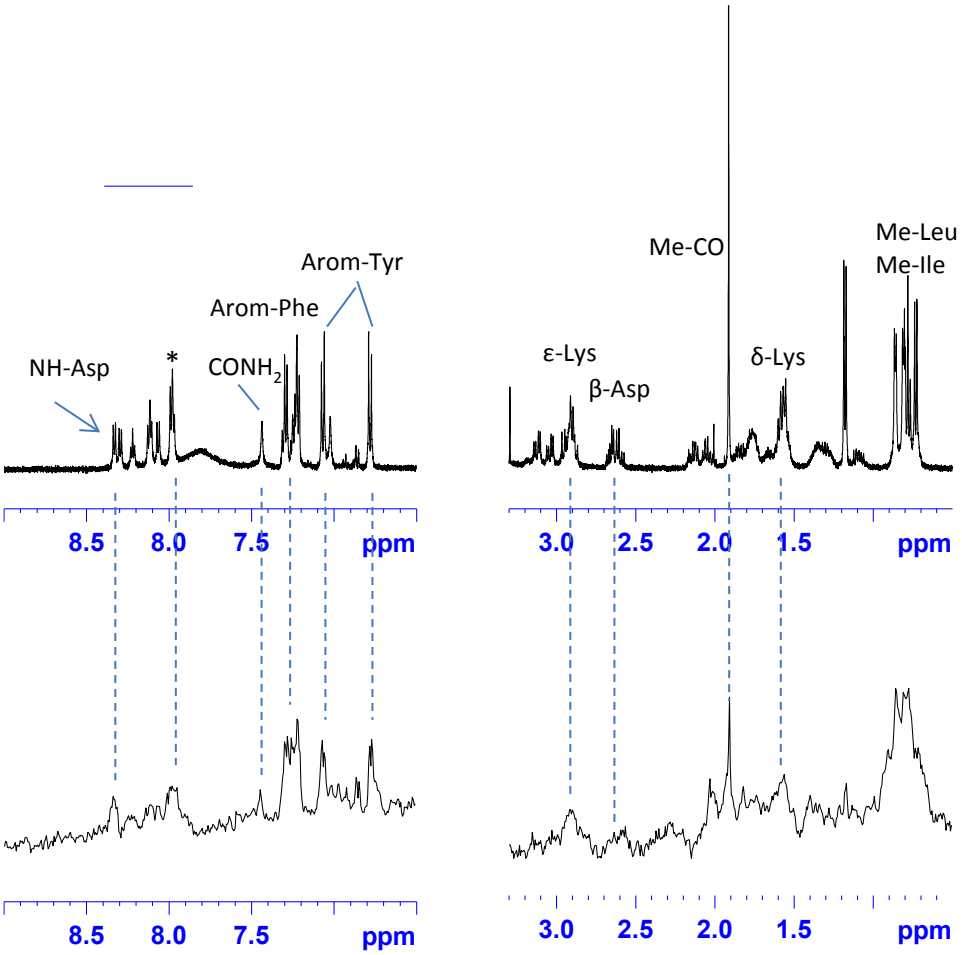

Fig. 3. Up) selected regions of ^{1}H-NMR spectrum of 1.1 mM peptide in the presence of 22 μM VP24 protein in phosphate buffer, resulting in a 50:1 ligand-target ratio. The signal labeled with an asterisk is due to the Lys, Leu and Ile amide protons which show the same chemical shift. Bottom) The same region of the STD spectrum.

To better assess the interaction strength, the percentage of absolute STD was calculated (Fig. S10, Table S2). Figure 4 shows the absolute STD percentage (grouped in two intensity ranges) for protons which forms the epitope that reflects the relative proximity of the atoms to the binding site. In fact, the ligand protons nearest to the protein are most likely to be saturated to the highest degree, and therefore have the strongest signal in the mono-dimensional STD spectrum. In particular, the more strictly interacting moieties are the aromatic protons of Tyr 477' and Phe 484' (1.5 and 1.6 % absolute STD respectively), the amidic proton of Asp 480' (2 % absolute STD) and the C-terminal amide protons (1.6 % absolute STD). Although CAS analysis and STD point out different aspect of peptide protein interaction (CAS is focused on the energetic of side chain interactions highlighted by their mutation into alanine, while STD supplies information on the proximity of a set of hydrogen atoms of the peptides with respect to the protein) we may notice that, by and large, the two approaches give mutually compatible and sound results.

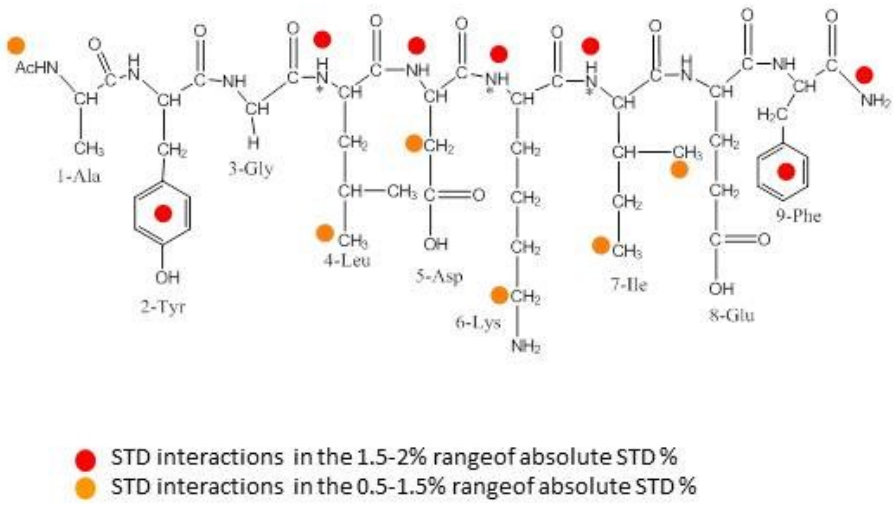

Fig. 4. Absolute STD% values, grouped in two intensity ranges, for peptide interacting with VP24 toxins. The NH-amides of Leu, Lys and Ile, labeled with an asterisk, show the same chemical shift.

The key role of the aromatic residues Tyr 477' and Phe 484', which take part in hydrophobic interactions with VP24, are highlighted both by CAS and by NMR analysis. Moreover, the role of charged residues, in particular Lys 481' and Asp 480', while somehow overestimated by CAS, as often reported in literature[56,57], has been observed by STD analysis. In order to perform a negative control, we analyzed the linear octapeptide A (LGYGFVNY) in the presence of VP24 protein by STD experiment. The A peptide has six residues in common with peptide RS, but despite this the interaction with the protein is not observed as showed in figure 2-B (no STD signals). The observation that RS can actually form a complex with VP24 in solution and the overall agreement between NMR and computational data, support our initial hypothesis that the selected peptide could bind to VP24 and suggest the plausibility of the proposed binding site.

Circular Dichroism

As a further confirmation of the interaction between RS and VP 24, far-UV circular dichroism (CD) experiments were performed. This technique is, indeed, particularly useful for studying protein folding and interactions.[59] In 50 mM PBS at pH 7.4, RS peptide lacked to adopt an ordered conformation, as it was almost random coil (Fig. 5, red line). In the same condition, VP24 is characterized by two negative bands at 210 and 225 nm (amide $n \rightarrow \pi^*$ transition), that are typical of α helix (Fig. 5, black line). On the other hand, the absence of a positive band below 200 nm (amide $\pi \rightarrow \pi^*$ transition) indicated that VP24 is present in solution as a mixture different conformations. VP24/RS interaction was then evaluated. A solution containing a mixture of RS and VP24 in 50 mM PBS was incubated at room temperature and then CD spectrum was recorded. A comparison between the spectrum of experimental mixture (Fig. 5, blue line) and the arithmetic sum of the

two separated components (Fig. 5, green line) highlighted that, actually, an interaction occurred. Furthermore, the decrease in intensity of the negative Cotton effects at 210 and 225 nm in the experimental mixture indicated a loss of the helical content of the protein, upon this interaction.

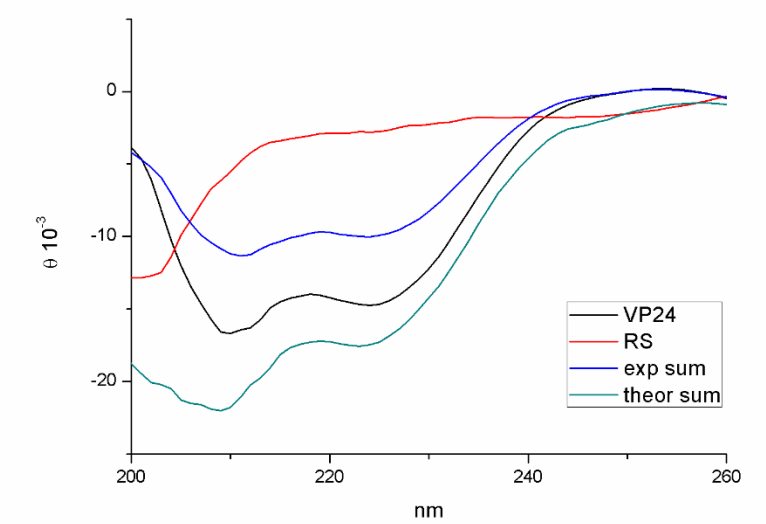

Fig. 5 CD spectra of VP24 (black line), of peptide RS (red line), and of VP24/RS experimental mixture (blue line). The green line is the arithmetic sum of VP24 and RS single spectra.

Conclusions

In this work we analyzed at the atomic level the network of interactions responsible for the formation of the VP24-KPNA complex. The protein-protein interface has been characterized, identifying the hot spots residues, that give a major contribution to binding energy. From this analysis, a nonapeptide ranging from Ala 476' to Phe 484' and comprising four hot spots close to one another both in the tertiary and primary structure ok KPNA has been identified. The ability of the selected peptide to actually interact with VP24 has been assayed with Saturation Transfer Difference NMR, which allowed to map the interaction epitope by determining the peptide regions in contact with the protein receptor. On the whole, the epitope sketched by NMR is consistent with the binding mode proposed by molecular modeling. The interaction between the selected RS peptide and VP24 was assessed also by far-UV CD experiments. The VP24 protein is a potential target for reducing Ebola virus virulence and lethality, but its structure, lacking well defined pockets and grooves is far from being easily druggable. The identification of the RS peptide, which interacts with VP24 through its surface, can be a useful starting point for further development of VP24 targeting active molecules.

Acknowledgmewnts

We thank Dr. Erica Ollmann Saphire (The Scripps Research Institute, La Jolla, California, U.S.A.) for kindly providing the plasmid for VP24 production. We acknowledge Fondazione Banca del Monte di Lombardia for financial support. Partial funding provided by Università degli Studi di Milano through the Development Plan of Athenaeum grant - line B1 was appreciated.

Bibliography

1. H. Feldmann and T. W. Geisbert, *The Lancet*, 2011, **377**, 849.
2. C. M. Bosio, M.J. Aman, C. Grogan, R. Hogan, G. Ruthel, D. Negley*et al*., *J. Infect. Dis.*, 2003, **188**, 1630.
3. M. Bray, T. W. Geisbert, *Int J Biochem Cell Biol*, 2005, **37**, 1560.
4. S. P. Reid, L. W. Leung, A. L. Hartman, O. Martinez, M. L. Shaw, C. Carbonelle*et al*., *J Virol*, 2006, **80**, 5156.
5. S. P. Reid, C. Valmas, O. Martinez, F. M. Sanchez, C. F. Basler, *J Virol*, 2007, **81**, 13469.
6. C.S. Brown *et al.*, *J. Mol. Biol.*, 2014, **426**, 2045.
7. P.S. Kharkar, P. Rasmasani, Y.S. Choong, L. Rhyman and S. Warrier, *RSC Adv.*, 2016, **6**, 26329.
8. F. Dapiaggi, S. Pieraccini and M. Sironi, *Mol. BioSyst.*, 2015, **11**, 2152.
9. S. Goodbourn, L. Didcock, and J. Randall, *J. Gen. Virol*, 2000, **81**, 2341.
10. X. Chen, U. Vinkemeier, Y. Zhao, D. Jeruzalmi, J.E. Darnell, and J. Kuriyan, *Cell*, 1998, **93**, 827.
11. K.M. McBride, G. Banninger, C. McDonald and N.C. Reich, *EMBO J.*, 2002, **21**, 1754.
12. W. Xu *et al*., *Cell Host & Microbe*, 2014, **16**, 187.
13. P. Chène, *ChemMedChem*, 2006, **1**, 400.
14. M.R. Arkin, *Nat. Rev. Drug. Discovery*, 2004,**3**, 301.
15. D.E. Scott, A.R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug. Discovery*, 2016,**15**, 533.
16. M. Raj, B.N. Bullock and P.S. Arora, *Bioorg. Med. Chem.*, 2013, **21**, 4051.
17. G. Zinzalla and D.E. Thurston, *Drugs of the Future*, 2009, **1**, 65.
18. R. Fasan, R.L.A. Dias, K. Moehle, O. Zerbe, J.W. Vrijbloed, D. Obrecht and J.A. Robinson, *Angew. Chem. Int. Ed.*, 2004, **43**, 2109
19. S. Pellegrino, N. Ferri, N. Colombo, E. Cremona, A. Corsini, R. Fanelli, M.L. Gelmi, C. Cabrele, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 6298.
20. J.A. Wells and C.L. McClendon, *Nature*, 2007, **450**, 1001.
21. M.R. Arkin *et al.*, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 1603.
22. T. Berg, *Angew. Chem. Int. Ed.*, 2003, **42**, 2462.
23. P.A. Kollman and I. Massova, *J. Am. Chem. Soc.*, 1999, **121**, 8133.
24. P.A. Kollman *et al.*, *Acc. Chem. Res.*, 2000, **33**, 889.
25. S. Pellegrino, L. Ronda, C. Annoni , A. Contini, E. Erba, M.L. Gelmi , R. Piano, G. Paredi, A. Mozzarelli and S. Bettati, *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2014, **1844**, 2108.
26. S. Pieraccini, G. Saladino, G. Cappelletti, D. Cartelli, P. Francescato, G. Speranza, P. Manitto and M. Sironi, *Nature Chemistry*, 2009, **1**, 642.
27. S. Pieraccini, S. Rendine, C. Jobichen, P. Domadia, J. Sivaraman, P. Francescato, G. Speranza and M. Sironi, *RSC Advances,* 2013, **3**, 1739.
28. D. Gonzales-Ruiz and H. Gohlke, *Curr. Med. Chem.*, 2006, **13**, 2607.
29. S. Chaurasia, S. Pieraccini, R. De Gonda, S. Conti and M. Sironi, *Chem. Phys. Lett.*, 2013, **587**, 68.
30. B. Meyer and T. Peters, *Angew. Chem. Int. Ed.*, 2003, **42**, 8, 864.
31. A. Bhunia, S. Bhattacharjya and S. Chatterjee, *Drug Discovery* Today, 2012, **17**, 505.
32. D. I. Freedberg and P. Selenko, *Annu. Rev. Biophys*. 2014, **43**, 171.
33. a) D. Potenza, F. Vasile , Ligand-target interaction: a NMR point of view. Referred book chapter in Seminars in Organic Synthesis. XXXVII A. Corbella Summer School, Società Chimica Italiana pp.224-248 - 2012. ISBN:978-88-86208-71-0. b) Guzzetti I, Civera M, Vasile F, Araldi EM, Belvisi L, Gennari C, Potenza D, Fanelli R, Piarulli U. *Org Biomol Chem*. 2013, **11**, 3886.
34. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4** , 435.
35. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950.

36. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
37. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaje, *J. Comput. Chem.,* 1997*,***18**, 1463.
38. G. Bussi, D. Donadio and M. Parrinello*, J. Chem. Phys.*, 2007, **126**, 014101.
39. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
40. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.,* 1993, **98**, 10089.
41. S. Hubbard and J. Thornton, Naccess, The University of Manchester, UK, 1992-6.
42. S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845.
43. N.A. Baker, D. Sept, S. Joseph, M.J. Holst, and J.A. McCammon, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 10037.
44. C. Paissoni, D. Spiliotopoulos, G. Musco and A. Spitaleri, *Comp. Phys. Comm.*, 2014, **185**, 2920.
45. D. Spiliotopoulos, A. Spitaleri and G. Musco, *PLoS ONE*, 2012, **7**, e46902.
46. S.A. Martins, M.A.S. Perez, I.S. Moreira, S. F Sousa, M.J. Ramos and P.A. Fernandes, *J. Chem. Theory Comput.*, 2013, **9**, 1311.
47. R.T. Bradshaw, B.H. Patel, E.W. Tate, R. L. Leatherbarrow and I.R. Gould, *PEDS*, 2011, **24**, 197.
48. A.P. Zhang, Z.A. Bornholdt, T. Liu, D.M. Abelson, D.E. Lee, S. Li, V.L.Woods and E.O. Saphire, *PLoS Pathog.*, 2012, **8**, e1002550.
49. A.P. Zhang, D.M. Abelson, Z.A. Bornholdt, T. Liu, V.L.Woods and E.O. Saphire, *Virulence*, 2012, **3**, 440.
50. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch, In The Proteomics Protocols Handbook (Walker JM ed) Humana Press, 2005, pp. 571-607.
51. S. Pellegrino, C. Annoni, A. Contini, F. Clerici, M.L. Gelmi, *Amino Acids*, 2012, **43**, 1995.
52. a) K. Wuthrich, *NMR of Proteins and Nucleic Acids*. New York (NY), USA: Wiley-Interscience. 1986. ISBN: 0471119172. b) F. Vasile, E. Pechkova and C. Nicolini, *Proteins*, 2008, **70**, 1112
53. I.S. Moreira, P.A. Fernandes and M.J. Ramos, *Theor. Chem. Acc.*, 2007, **117**, 99.
54. a) D.Potenza, L. Belvisi, F. Vasile, E. Moroni , F. Cossu and P. Seneci, *Org. Biomol. Chem.* 2012, **10,** 3278. b) Manzoni L, Belvisi L, Bianchi A, Conti A, Drago C, de Matteo M, Ferrante L, Mastrangelo E, Perego P, Potenza D, Scolastico C, Servida F, Timpano G, Vasile F, Rizzo V, Seneci P. *Bioorg Med Chem*. 2012, **20**, 6687.
55. a) F. Vasile, D. Rossi, S. Collina and D. Potenza, *EurJOC,* 2014, doi: 10.1002/ejoc.201403014. b) Heggelund JE, Haugen E, Lygren B, Mackenzie A, Holmner A, Vasile F, Reina JJ, Bernardi A, Krengel U. *Biochem Biophys Res Commun.* **418,** 731.
56. S. Pieraccini, R. De Gonda and M. Sironi, *Chem. Phys. Lett.*, 2011, **517**, 217.
57. S. Huo, I. Massova and P.A. Kollman, *J. Comput. Chem.*, 2002, **23**, 15.
58. W. Humpherey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33.
59. S. M. Kelly, T. J. Jess, N. C. Price BBA, 2005, 1751, 119.