Crystal structures of LCMV endonuclease domain complexed with diketo acid ligands.

Magali Saez-Ayala[2,3], Elsie Laban Yekwa[1,2,4],Mauro Carcelli[5], Bruno Canard[1,2], Karine Alvarez[1,2]*, François Ferron[1,2]*

1 CNRS, AFMB UMR 7257, 13288, Marseille, France

2 Aix-Marseille Université, AFMB UMR 7257, 13288, Marseille, France

3 current address : Aix-Marseille Université, CRCM, INSERM U1068, CNRS UMR7258 , 13273, Marseille, France

4 current address : Division of Medical Virology, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa

5 Dipartimento di Scienze Chimiche, della Vita, della Sostenibilità Ambientale P.co Area delle Scienze 17/A 43124 Parma, Italy

**Contributed equally. Corresponding author: francois.ferron@afmb.univ-mrs.fr*

Synopsis - Crystallography applied to molecule optimization to its molecular target

Abstract - The *Arenaviridae* family is one of the three negative stranded RNA viral families encoding an endonuclease in their genome, together with *Bunyaviridae* and *Orthomyxoviridae*. The endonuclease domain is carried at the N-terminal of the L protein, a multifunctional protein that includes the RNA dependent RNA polymerase. The synthesis of mRNA in arenaviruses is a process that is primed by capped nucleotides stolen from the cellular mRNA by the endonuclease domain in cooperation with other domains of L protein. This molecular mechanism had been demonstrated earlier by our group on the endonuclease of the prototype Lymphocytic ChorioMeningitis virus (LCMV). However, the mode of action of this enzyme is not fully understood as the original structure did not contain the catalytic metal ions. The pivotal role played by the cap-snatching process in the life cycle of the virus and the highly conserved nature of the endonuclease domain makes it a target of choice for the development of novel antiviral therapy. Here we evaluated using biophysical methods the binding affinity of two diketo acids (DKAs) compounds (DPBA (**1**) and L-742,001 (**2**)) onto the LCMV endonuclease domain. We have determined the X-ray structures of LCMV endonuclease domain with catalytic ions in complex with these two compounds, and assessed their efficacy in an *in vitro* endonuclease activity assay. Based on these data and

computational simulation, we synthesized two new DKAs. The LCMV endonuclease exhibits a good affinity for these DKAs making them a good starting point for the design of arenavirus endonuclease inhibitors. Beside being the first example of an arenavirus endonuclease X-ray structure incorporating a ligand, this study is a proof of concept that the design of optimized inhibitors against the arenavirus endonuclease is possible.

Keywords: - Arenaviridae ; endonuclease ; LCMV ; diketo acids; compounds optimization; metal chelation.

1. Introduction

Arenaviridae is a family of viruses associated with rodent transmitted infections in humans. These viruses cause chronic and asymptomatic infections in rodents, and constitute a reservoir of human pathogens across the world (Buchmeier *et al.*, 2007). Indeed, several *Arenaviruses* including Lassa, Lujo, Gunarito, Sabia, Chapare, Machupo, and Junin viruses are responsible for spiking hemorrhagic fever outbreaks constituting a major public health concern (Bowen *et al.*, 1997; Enria *et al.*, 2008; Briese *et al.*, 2009). In West Africa alone (Sierra Leone, Guinea, Liberia, Nigeria, Benin), Lassa virus is responsible for several hundred thousand infections per year (Günther & Lenz, 2004). It is a common endemic infection that in most cases will lead to hearing loss, tremors and encephalitis, and for 1% of cases will turn to a deadly hemorrhagic fever (Yun *et al.*, 2015). The last Lassa fever outbreak between August 2015 and May 2016 in Nigeria reported a case fatality rate (CFR) of 53.9% based on confirmed WHO data of 165 cases and 89 deaths (confirmed through laboratory testing) (W.H.O, 2016). Lymphocytic choriomeningitis virus (LCMV) was the first arenavirus to be isolated and is considered as the prototypic virus for the family. Responsible for occasional transmission to man, it may result in life-threatening meningitis and/or hemorrhagic fever, and several clinical studies suggest that its pathogenicity is underestimated (Mets *et al.*, 2000; Schulte *et al.*, 2006; Jamieson *et al.*, 2006; Bonthius, 2012). Because its natural host is the common house mouse (*Mus musculus*), it has a global distribution with LCMV infections reported in Europe, the Americas, Australia, and Japan. LCMV is a human pathogen of significant clinical relevance, causing central nervous system disease, congenital malformation, choriomeningitis, and systemic and highly fatal infection in immuno-compromised, organ transplant recipient patients (Fisher-Hoch *et al.*, 1995; Mets *et al.*, 2000; Barton *et al.*, 2002; Fischer *et al.*, 2006; Schulte *et al.*, 2006; MacNeil *et al.*, 2012). Humans are generally infected by arenaviruses through exposure to aerosols of fresh urine, droppings, saliva, nesting materials from infected rodents, or during organ transplantation (Fischer *et al.*, 2006). Despite the widespread threat to human health that these pathogens represent, there are no vaccines and only limited therapeutic options (Lee *et al.*, 2011). The only licensed drug for the treatment of human arenavirus infection is the broad-spectrum antiviral ribavirin, which is

only partially effective and is associated with significant toxicities (McCormick *et al.*, 1986; Kilgore *et al.*, 1997; Enria *et al.*, 2008). Therefore, there is an unmet need to identify novel compounds that could be developed into antiviral drugs to combat human-pathogenic arenaviruses.

Arenaviruses are enveloped viruses with a bi-segmented negative single-strand RNA genome. Each segment encodes in an ambisense manner two proteins separated by an intergenic region (IGR). The L RNA segment encodes a large protein L (~200 kDa) and a small disordered protein Z (~ 11 kDa) (Buchmeier *et al.*, 2007). L (P14240) is a multi-domain protein including in its N-terminus an endonuclease domain (residues 1 - 196) followed by a viral RNA-dependent RNA polymerase (RdRp) domain but lacks a capping machinery (López *et al.*, 2001). The Z protein, which contains a RING finger motif, is a multifunction protein regulating the life cycle of the virus and during budding assembles to form the matrix (Cornu & de la Torre, 2001; Perez *et al.*, 2003, 2004; Hastie, Zandonatti *et al.*, 2016). The S RNA segment encodes the precursor of mature virion glycoprotein GP-C (75 kDa); that will give after post-translational cleavage GP-1 (40 to 46 kDa) and GP-2 (35 kDa) (Hastie, Igonet *et al.*, 2016) ; and nucleoprotein NP (~ 63 kDa) (Buchmeier *et al.*, 2007). NP forms a polymer protecting the genomic (and anti-genomic) RNA (RNA_v). L and NP together with RNA_v form an active ribonucleic complex for replication and transcription (Pinschewer *et al.*, 2003). Arenaviruses belongs to negative RNA viruses, therefore the first step of the viral cycle starts by the transcription of L and NP viral messenger RNAs (mRNA) (Perez & de la Torre, 2003). This step is undertaken by the L protein and starts (maybe with the help of NP) to recruits cellular mRNA as first step. The L protein through its endonuclease domain cleaves and snatches the native cellular mRNA cap structure after ~11 nucleotides, and uses it as primer sequence for the L RdRp domain to synthesize the viral mRNA. The original crystal structure of the LCMV endonuclease domain (ENDO) was determined some time ago (Morin *et al.*, 2010), but in this original structure were missing the catalytic ions necessary to fully understand the mode of action of the endonuclease. Cap-snatching is essential to the life cycle of the virus and it makes the endonuclease domain a target of choice for drug design development (Charrel *et al.*, 2011). This mechanism is shared by *Orthomyxoviridae* and *Bunyaviridae;* two families harboring prominent pathogen members; and for which the endonuclease was used in drug development (Reguera *et al.*, 2010; Kowalinski *et al.*, 2012; DuBois *et al,* 2012). However, *Arenaviridae* endonuclease is a challenging target for inhibitor development. Indeed, it has the most phylogenetically remote structure compared to the enzymes from *Orthomyxoviridae* or *Bunyaviridae* (Ferron *et al.*, 2017). Previous structural studies on Lassa Virus endonuclease characterized the complex with catalytic ions (Wallat *et al.*, 2014) but failed to obtain a complex with an inhibitor (Reguera et al., 2016).

Nevertheless, there is large interest in viral endonuclease inhibitors that are optimally designed to occupy the active site and chelate the metal ions (Rogolino *et al.*, 2012).

In this study we selected two molecules that are known to target the Influenza virus endonuclease (*Orthomyxoviridae*) as a starting point for the structure-based-drug-design of anti-arenavirus compounds (Noble *et al.*, 2012; Stevaert *et al.*, 2013). We characterized by biophysical methods the binding affinity of the DPBA (**1**) and L-742,001 (**2**) (Fig. 1a) to the LCMV ENDO. We evaluated their efficacy in an *in vitro* endonuclease assay and we solved the crystal structures of LCMV ENDO with catalytic ions (Mg^{2+} or Mn^{2+}), and each of these two diketo acids (DKAs). Based on the structure analysis and *in silico* analysis, we synthesized two new DKAs, compounds **3** and **4** (Fig. 1b), which exhibited not only a better affinity but also a better inhibition regarding the LCMV ENDO activity.

2. Materials and methods

2.1. Sequence analysis

All available full length sequences of Arenaviridae L protein where downloaded from NCBI. Using the Jalview redundancy option all identical sequences were removed. The working set was composed of 384 sequences including Mammarenavirus and Reptiarenavirus sequences. That subset was aligned with Muscle (Edgar, 2004) with the balanced option optimized for long sequence and large data set. The first 220 amino acids corresponding to an extended endonuclease domain were selected. The working set was purged once more of identical sequences leaving 245 sequences. The resulting alignment was analysed to target specifically identical residues and the information was correlated to LCMV endonuclease structure. The motifs conservation was represented with Web Logo (Crooks *et al.*, 2004) (List of sequences used is summarized in Table S1 and alignment done with ESPript ((Gouet *et al.*, 2003) in Fig. S1).

2.2. Protein expression and purification

The ENDO wild type (WT) and D118A and D88A mutants were cloned into pDest14 with a N-terminus hexa-histidine tag and expressed in E.coli Rosetta (DE3) pLysS, at 17ºC in TB medium overnight after induction with 500 mM IPTG. Cell pellets from harvested cultures were resuspended in 50 mM TRIS buffer, pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, and 5% Glycerol. Lysozyme (0.25 mg/ml), DNase I (10 µg/ml), and EDTA free protease cocktail (Roche) were added before sonication. IMAC chromatography of clarified lysates was performed on a 5 ml His prep column (Akta Xpress FPLC system, GE Healthcare) eluted with the same buffer with 500mM imidazole. The eluted his-tagged fraction was diluted and purified on HiTrap Q sepharose 1 ml column (GE Healthcare).

Proteins were eluted in a linear gradient from 50 mM to 1 M NaCl in 10 mM HEPES buffer, pH 7.5, and 2 mM DTT. Size exclusion chromatography was performed on preparative Superdex 200 column (GE Healthcare) pre-equilibrated in 10 mM HEPES, pH 8.0, 50 mM NaCl, and 2 mM DTT. Protein was concentrated to 25 mg/ml and frozen in liquid nitrogen.

2.3. Compounds and substrate characterization

2.3.1. Differential Scanning Fluorimetry (DSF)

Melting temperature (T_m) values of the proteins were determined by a thermo-fluorescence based assay. In 96-well thin-wall PCR plates, 11 µl of protein (ENDO) were added to 11 µl of compound (**1** or **2**) and/or metal ions (Mg^{2+} or Mn^{2+}), in 10 mM HEPES buffer, pH 8.0, 50 mM NaCl, and 2 mM DTT. Finally, 3 µl of the fluorescent dye Sypro Orange was added (Molecular Probes, 715-fold diluted in H_2O). Thermal denaturation of the proteins was followed by measuring fluorescence emission at 575 nm (with excitation at 490 nm) in a CFX Connect Real-Time PCR Detection System (Biorad). Final concentrations were adjusted to 75 µM of protein, 0.5 mM of $MgCl_2$, 0.5 mM of $MnCl_2$, 450 µM of compound (**1** or **2**) (final ligand/protein ratio = 6), and 5% DMSO. Denaturation midpoints of proteins were calculated using the Boltzmann equation using GraphPad Prism. All measurements were performed in triplicates.

2.3.2. Microscale thermophoresis (MST)

MST experiments were performed on a Monolith NT.115 instrument (NanoTemper Technologies). Proteins were labelled with the red fluorescent dye NT-647 using the Protein Labeling Kit RED-NHS (NanoTemper Technologies). The concentration of labelled protein was kept constant at 100 nM, while the concentration of compound was varied. A 15-step two-fold dilution series beginning from 500, 250 or 125 µM, finally yielded 16 different concentrations of tested compound (**1** to **4**). Experiments were carried out in 10 mM HEPES buffer, pH 8, containing 100 mM NaCl, 1 mM DTT, 0.05% (w/v) Tween-20, 0.25 mM $MgCl_2$ and 0.25 mM $MnCl_2$. Final samples were adjusted to 5% DMSO to ensure solubility of compounds. The samples were centrifuged for 5 minutes at 13000 rpm to remove potential aggregates, and the supernatant was loaded into standard treated MST-grade glass capillaries (NanoTemper Technologies). After 5 min incubation period the MST was measured with 80% LED power and 80% infra-red laser power. K_D values were determined using the NanoTemper analysis software.

2.3.3. Isothermal titration calorimetry (ITC)

Purified ENDO was diluted in the ITC buffer containing 10 mM HEPES, pH 8.0, 50 mM NaCl, 2 mM DTT, 0.25 mM $MgCl_2$ and 0.25 mM $MnCl_2$. Compound **1** was diluted in the

same buffer. Final DMSO concentration was 5% in cell and syringe. Titrations were carried out on a MicroCal ITC200 microcalorimeter (GE Healthcare). Experiments were designed using a titrant concentration (1mM compound DPBA (**1**) in the syringe) 10 times the analyte concentration (100 µM protein in the cell), using 19 injections at 25°C. A first small injection (0.2 µL) was included in the titration protocol in order to remove air bubbles trapped in the syringe prior titration. Raw data were scaled after setting the zero to the titration saturation heat value. Integrated raw ITC data were fitted to a one site non-linear least squares fit model using the MicroCal Origin plugin as implemented in Origin 9.1 (Origin Lab). Each experiment was performed three times and the data are presented as the mean ± S.D.

2.3.4. *In vitro* endonuclease assay

An *in vitro* endonuclease assay was performed to investigate the inhibition of compounds **1** and **2**. The reaction was carried in 20 µl samples containing 20 mM TRIS-HCl pH 8, 150 mM NaCl, 1 mM TCEP and 2 mM $MnCl_2$. ENDO-WT (20 µM) was premixed with 50, 125, 250, 500, 1000 or 2000 µM of each compound and the reaction was started by adding 1 µM of a 19 nucleotide single stranded RNA (Morin *et al.*, 2010). The reaction mixtures were incubated at 25°C and the reaction stopped after 6 h by adding 20 µl of loading buffer (formamide containing 10 mM EDTA). For experiments to compare the inhibition of DPBA (**1**) and compounds **3** and **4** , the ENDO-WT (20 µM) was premixed with 50 µM of each compound and the same protocol as above was applied. The reaction products were analyzed in 20% polyacrylamide / 8 M urea gels. The undigested RNAs were visualized using photo-stimulated plates with the Fluorescent Image Analyzer FLA3000 (Fuji) and, quantified using Image Gauge (Fuji). Graph were plotted using GraphPad PRISM. All experiments were performed in triplicates.

2.4. Crystallization, data collection and structure determination.

2.4.1. Crystallization

All crystals grew at 20°C using the sitting drop vapor diffusion method in 96 well plates (Greiner) by mixing 200 nl of protein solution (13 mg/ml in 10 mM HEPES, pH 8.0, 50 mM NaCl, 2 mM DTT) with 100 nl of reservoir solution using Mosquito robot (TTP Labtech). Crystals of ENDO-WT complexed with ions grew in a reservoir solution containing 100 mM sodium citrate, pH 6,2 and 3% Isopropanol, using a protein solution supplemented with 1mM $MgCl_2$. Crystals of ENDO-WT complexed with ions and DPBA (**1**) grew in a reservoir solution containing 100 mM sodium citrate, pH 6,2 and 3.5% Isopropanol, using a protein solution supplemented with 1mM $MgCl_2$ and 2mM DPBA (**1**) (solubilized in PEG 400). Crystals of ENDO-WT complexed with ions and L-742,001 (**2**) grew in a reservoir solution

containing 100 mM sodium citrate, pH 6 and 3.5% Isopropanol, using a protein solution supplemented with 1mM $MnCl_2$ and 2mM L-742,001 (**2**) (solubilized in PEG 400). Crystals of ENDO-D118A mutant grew in a reservoir solution containing 100 mM sodium citrate, pH 6,3 and 8% Isopropanol. All crystals were cryo-cooled in liquid N_2 using 20% v/v glycerol in mother liquor as cryoprotectant.

2.4.2. Data Collection and Structure Determination LCMV

The data of the ENDO-WT complexed with ions were collected on ID23-1 and ENDO-WT complexed with L-742,001 (**2**) data were collected on ID23-2 at the European Synchrotron Radiation Facility (ESRF). ENDO-WT complexed with DPBA (**1**) and the ENDO-D118A mutant data were collected on Proxima 1 at Synchrotron Soleil. Data were processed, analysed and scaled using the autoProc toolbox (Vonrhein *et al.*, 2011). Phases were obtained by molecular replacement using Phaser (McCoy *et al.*, 2007) and PDB: 3jsb (Morin *et al.*, 2010) used as search model. All water and ligand molecules were removed from the search structure. Rebuilding of the initial model using AutoBuild (Terwilliger *et al.*, 2008) was then performed. Subsequently, all-atom isotropic temperature-factor refinement cycles were performed with phenix.refine (Afonine *et al.*, 2012) and or with Buster (Blanc *et al.*, 2004). Electron-density maps were inspected using Coot (Emsley *et al.*, 2010). Extra density accounting for ions, and or compounds was observed for the complexed structures but none was found for mutant structures (supplementary Fig. 4B). The structures were evaluated using molprobity (Chen *et al.*, 2010) and PROCHECK (Laskowski *et al.*, 1993). Structural analysis and high resolution figures were done with UCSF chimera (Pettersen *et al.*, 2004*b*). Statistics of data collections and refinements are given in Table 1.

2.5. Molecular docking

For the molecular docking we used the crystal structures of ENDO-WT complexes with different ions. The 3D structure of ENDO-WT was energy minimized by the steepest gradient method of energy minimization followed by conjugate gradient minimization, using the MMTK and Amber packages (Cornell *et al.*, 1995; Hinsen, 2000; Lindorff-Larsen *et al.*, 2010). Compounds were designed using Bkchem (Kosata & Danne, 2010) and geometry restrained using eLBOW (Moriarty *et al.*, 2009). Mol2 and PDB files format of the ligands and receptor were converted to PDBQT format using UCSF chimera prior to docking. All the water and solvent atoms of the protein were removed and the polar hydrogen and polar charge were added onto the ions and ligand prior to docking. The protein was kept rigid while the ligand was allowed to rotate and explore more flexible binding pockets. Docking of the ligands onto the ENDO-WT was performed iteratively using AutoDock vina - version 1.1.2 (Trott & Olson, 2010). The best poses from the first round of docking were used as seed for

the second round. The grid box size dimensions were first 40X40X40, to verify that our ligands will preferentially bind in the catalytic site. The grid box size was further optimized to 23.2X15.6X21.2 thus covering the binding pockets, the default scoring function was used for docking.

Binding modes of the docked complexes were obtained and sorted based on their binding energy, ions and amino acid residues present at a distance less of 3 Å were considered as the binding partners of the ligands. Binding modes were compared to these of the native structure. The interaction figures representing the docked complexes have been generated using chimera (Pettersen *et al.*, 2004*a*).

2.6. Chemicals and Chemistry

2,4-dioxo-4-phenylbutanoic acid (DPBA **1**) was purchased from Interchim. (Z)-4-[1-benzyl-4-[(4-chlorophenyl)methyl]piperidin-4-yl]-2-hydroxy-4-oxobut-2-enoic acid (L-742,001 **2**) was synthesized as previously reported (Stevaert *et al.*, 2015). 2-Hydroxy-4-(biphenyl-4-yl)-4-oxobut-2-enoic acid (**3**) and 2-Hydroxy-4-oxo-4-(phenanthren-3-yl)but-2-enoic acid (**4**) were synthesized as described in literature (Patil *et al.*, 2007; Bhatt *et al.*, 2011) and conditions are detailed in Appendix A.1.

3. Results and discussion

Endonuclease catalytic reaction is driven by a two-metal ion catalysis (TMIC): 1) to form a stable intermediate, and then 2) to break the phosphodiester backbone of the nucleic acid chain (Palermo *et al.*, 2015). Indeed, as shown by Steitz et al. (Steitz & Steitz, 1993) the TMIC reaction is driven by two divalent metal ions positioned on each side of the targeted scissile phosphate bond to be cleaved. The metal ions facilitate an S_N2-type reaction, in which one metal ion favours the formation of the nucleophilic oxygen from a water molecule and the other helps the exit of the leaving group by associating with a non bridging phosphoryl oxygen. The catalytic center of the ENDO resides in the N-terminal part around the residue D88 and should contain two Mg^{2+} (or Mn^{2+}) ions, which are critical for the RNA cleavage mechanism. Our strategy to design inhibitors of this activity is to develop compounds able to block and inhibit the catalytic center through the trapping of critical metal ions, resulting in a functional impairment.

DPBA (**1**) and L-742,001 (**2**) are DKAs carrying a chelating motif, consisting of a γ-ketone, enolizable α-ketone and carboxylic acid, able to chelate divalent metal ions present in the active site of some enzymes (Kowalinski *et al.*, 2012; DuBois *et al.*, 2012) (Fig. 1). These compounds have been identified as potent inhibitors of the influenza PA N_{ter} activity in enzymatic assays or cell cultures and mouse models (Reguera *et al.*, 2010; Carcelli *et al.*,

2014; Stevaert *et al.*, 2015). This prompted us to investigate if they can also have a similar effect on LCMV ENDO which is an homologue of influenza PA N_{ter}.

3.1. Biophysical characterization of DKA compounds onto ENDO

3.1.1. Thermal stability by DSF

In order to characterize the effect of DKAs **1** and **2** on the endonuclease domain, we first investigated the thermal stability of ENDO-WT and two mutants; ENDO-D88A and ENDO-D118A. The two ENDO mutants (D88A and D118A), were selected to elucidate the role of the two Aspartates in metal and ligand binding. The D88A is a mutation on the key residue supposedly involved in bivalent metal ion coordination, while the D118A is a mutation away from the active site of a conserved residue in the vicinity of the catalytic pocket (Morin *et al.*, 2010) . To determine the thermal stability of ENDO-WT and its mutants, we used a thermo-fluorescence based assay. In this assay, the fluorescence of a hydrophobic dye is measured as it binds to hydrophobic regions of the protein that are solvent-exposed as the protein is denatured by heating. The melting temperature (T_m) of the protein is then determined from a temperature dependent fluorescence curve. As shown in Fig. 2a, changes in T_m of ENDO-WT and the ENDO-D118A mutant are behaving in the similarly contrary to the ENDO-D88A mutant.

In the absence of ions, the T_m of both ENDO-WT and ENDO-D118A mutant are around 40ºC, whereas that of the ENDO-D88A mutant is 10° C higher indicating that the enzyme is more stable. In the presence of ions, for both ENDO-(WT or D118A), the addition of divalent ions (Mg^{2+}, Mn^{2+} or mixture of both) increases the T_m by ~5° C. On the other hand, ENDO-D88A, in similar conditions, presents no significant T_m increase. These results show that: *i*) ENDO-(WT or D118A) are stabilized by divalent ions but not ENDO-D88A. *ii*) The residue D118 has no direct or indirect interaction with the ions, and has virtually no influence on the protein stability. This result was confirmed by our structural study (see paragraph 3.2). *iii*) The mutant D88A is not affected by either the presence or absence of divalent ions, indicating that residue D88 is involved in divalent ion binding.

We next assessed the effect of divalent metal ions and ligand binding onto ENDO-WT and mutants (Fig. 2a). For both ENDO-(WT or D118A), the addition of either DPBA (**1**) or L-742,001 (**2**) in the presence of metal ions led to a significant positive T_m shift of more than 10° C, which confirmed the high affinity of these ligands for the enzyme. No effect was observed when using the ENDO-D88A. This latter result provides evidence that in absence of divalent ion, the enzyme displays no affinity for the compounds. The increase of $\mathbf{T_m}$ indicates a gain of stability due to the binding effect of the compounds onto ENDO complexed with divalent ions. As no gain of stability was observed with the D88A mutant, a mutation that

abrogates metal ion binding, we infer that the mode of binding of ligands **1** and **2** is done through metal ion chelation.

3.1.2. Binding affinity by MST and ITC

To further characterize the ligand binding mode, we used Microscale Thermophoresis to determine the affinity of ENDO-WT and mutants ENDO-D88A and ENDO-D118A for DPBA (**1**) and L-742,001 (**2**), in the presence of divalent metal ions (Mg^{2+} and Mn^{2+}). MST detects changes in the hydration shell of biomolecules and measures their interactions under close-to-native conditions. Any change of the hydration shell of proteins due to changes in their structure affects the thermophoretic movement, which is used to determine binding affinities with high accuracy and sensitivity. Results summarized in Table 2 show that the affinity of the ENDO-WT and ENDO-D118A for DPBA (**1**) are in the same micromolar range. However, in the case of L-742,001 (**2**), the K_D values for ENDO-WT and ENDO-D118A were about 10 and 20 fold lower than that for DPBA (**1**) respectively. This suggests that both ENDO-WT and ENDO-D118A display higher affinity for L-742,001 (**2**) than for DPBA (**1**).

Interestingly, for both compounds**,** no significant affinity is recorded for either ENDO-WT without metal ions in the experimental buffer or for ENDO-D88A. This result corroborates the idea that the binding of the DKAs **1** and **2** takes place *via* divalent cations chelation and that residue D88 is key in this interaction.

We further validated results obtained with DPBA (**1**) using Isothermal Titration Calorimetry (ITC) as an orthogonal assay. The K_D of DPBA (**1**) to ENDO-WT and the two mutants, ENDO-D118A and ENDO-D88A, were determined and data were of the same order to those determined by MST (Fig. S2). ENDO-WT and ENDO-D118A displayed affinity in the low micromolar range and for ENDO-D88A, no interaction was detected, confirming the earlier results.

3.2. Holoenzyme ENDO-WT and ENDO-D118A structures

In the original study of LCMV ENDO (Morin *et al.*, 2010), crystallization conditions did not allow the formation of complexes with co-factors (ionic or other wise). This was in part due to the organization of the molecules in the crystals asymmetric unit, one symmetric molecule occupying the catalytic site. Therefore for this structural study to be successful it was critical to search for new crystalization conditions that would allow the molecules to form ionic complexes. The present ENDO-WT structure clearly shows the two Mg^{2+} ions directly coordinated by the catalytic Aspartate D88 and through a network of water molecules, by Glutamic acid E101 and Lysine K114 (Fig. S3a). From our observation of 100

crystals during screening and the analysis of work performed on Lassa virus endonuclease domain, this event of coordination seems to be rare. Indeed, in the structure of Lassa Virus endonuclease domain of Wallat *et al.*, (2014) similar residues are involved with a very dense network of water molecules stabilizing the ions. On the other hand, if we compare these results to the structure of Lassa Virus endonuclease domain of Reguera *et al.*, (2016), we can see that the two Mn^{2+} coordination involved the corresponding residues D89, E102 and K115 and a residue from the asymmetric unit (E3) and no water. Altogether these data form a trend indicating that the affinity of the *Arenaviridae* endonuclease for its catalytic ions is low and the metal ion binding at the catalytic site is the result of a random event in the absence of a RNA substrate.

In the ENDO-WT structure, the water molecules mimic the positions of the non-bridged oxygens of the RNA phosphodiester backbone involved in the binding with the ions. This observation reminded us of previous comments on Ribonuclease III; for which the substrate is proposed to bind to the enzyme independently and prior to the catalytic ions (Nicholson, 2014). Likewise, this is also the case for the *Arenaviridae* exonuclease mechanism for which the RNA contributes to the formation of the binding site of the catalytic ion (Jiang *et al.*, 2013).

Therefore, we propose that in the case of the LCMV endonuclease domain, the RNA substrate is a necessary contributor to the formation of a transitioning ion binding/catalytic site that allows a two metal ion catalytic mechanism (TMIC).

In this study, we decided to assess the impact of a key residue within the catalytic site (Morin *et al.*, 2010) but *a priori* not involved in the mechanism. For this, we mutated the conserved Aspartate into Alanine (D118A). From a structural point of view, we confirmed by crystallography that the mutation has no effect on the overall LCMV endonuclease structure (Fig. S3b), as preliminary shown in the study of the thermal stability of the ENDO-D118A by DSF.

3.3. DPBA-ENDO-WT complex

To understand the molecular interactions between the DPBA (**1**) and ENDO-WT, we solved their co-structure. The DPBA-ENDO-WT complex was determined at 1,88 Å resolution in the space group P41. DPBA (**1**) binds directly to the two Mg^{2+} ion in the active site as clearly shown by the 2Fo-Fc omit map (Fig. 3a, b, and c). However, this binding mode is rather unusual compared to the binding configuration observed for DPBA (**1**) binding to the active site of Ortho-bunyavirus endonuclease and the influenza PA N_{ter} (Reguera *et al.*, 2010; Kowalinski *et al.*, 2012; DuBois *et al.*, 2012). Indeed, DPBA (**1**) binds the first ion through the carboxylic group and the enolizable α-ketone and the second ion through both α/γ-ketones. In DPBA-ENDO-WT structure, the three adjacent and co-planar oxygen atoms of the

DKA motif of DPBA (**1**) bind the two metal ions as follows: the Mg1 is coordinated by the carboxylic group and enolizable α-ketone, whereas the Mg2 is coordinated by the carboxyl group only (Fig. 3a, b and S4a). The γ-ketone is flipped by about 180° and forms hydrogen bond with two water molecules mediating interaction with residues S46 and E50. The phenyl group of DPBA (**1**) makes no direct interactions with residues in the active site.

This difference in binding mode might be due to the enlarged nature of the active site of the LCMV ENDO versus Orthobunya- or influenza endonucleases, allowing the compound to adopt alternative conformations.

3.4. L-742,001-ENDO-WT complex

The L-742,001-ENDO-WT complex was determined in the same space group as the DPBA -ENDO-WT complex P41, at 1,97 Å. Contrary to DPBA (**1**), L-742,001 (**2**) adopts a conformation that is identical to the one reported for influenza PA N_{ter} (Kowalinski et al., 2012; DuBois et al., 2012). The two metal ions are chelated by the three co-planar oxygen atoms of L-742,001 (**2**) as indicated on the 2Fo-Fc omit map (Fig. 3d, d f and S4b). One of the Mn^{2+} ions is coordinated by D88, E50 as well as D65 and K121 via bridging water molecules and to two oxygen atoms of the ligand. The second metal ion is coordinated by C102, D88, two water molecules and two oxygen atoms of the ligand, leading to an octahedral geometry for both ions. The oxygen atom of the ligand also makes a hydrogen bond with K114, an important catalytic residue. The chlorobenzyl and benzyl-piperidine groups are oriented in opposite directions perpendicular to the dioxo-butanoic acid. The chlorobenzyl group enters into a pocket comprising R47, S46, K43 and E50. The piperidine moiety on the other hand is poorly defined due to weak electron density as compared to the rest of the molecule. This suggests that there exist some rotational flexibility due to weak interactions between this part of the molecule and residues in the flexible region (residues 83 to 85) connecting the fourth α-helix to the first β-strand of the protein. Moreover, we observed an alternative conformation for the side chain of D88, suggesting that the presence of the compound taking away the ion was enough to alter the “resting” position of the side chain.

The overall 3 structure comparison (holo- and 2 ligands complexed) raise an interesting observation concerning the way the ENDO coordinates the ions, depending on the active site environment. Indeed in the holo- structure, ions are coordinated by three residues D88, E101 and K114; in DPBA (**1**) complexed structure ions are also coordinated by three residues D88, E101 and D118 (via a water molecule). On the other hand, in L-742,001 (**2**) complexed structure ions are coordinated by E50, D88 and K121. The obvious difference between the two ligands lays in the fact that DPBA has no direct interaction with the enzyme, while L-742,001 (**2**) does interact through its chlorobenzyl group. This observation shows that any

perturbation in the active site, such as a ligand, can trigger ions mobility. Figuratively, the ions bounce like a pinball in the active site to establish a dynamic new coordination.

3.5. Evaluation of DPBA (1) and L-742,001 (2) in an *in vitro* LCMV endonuclease activity assay

Having characterized the interactions between the two ligands and ENDO-WT, we proceeded to test the efficacy of compounds DPBA (**1**) and L-742,001 (**2**) to inhibit the LCMV endonuclease activity in an *in vitro* endonuclease assay. A single stranded 5' radio-labeled RNA was incubated with the ENDO-WT with a range of compound concentrations. Cleavage was analysed on a denaturing polyacrylamide gel and visualized on a phosphoimager (Fig. 2b). As expected, DPBA (**1**) and L-742,001 (**2**) display inhibition of LCMV endonuclease activity, however, both compounds exhibit weak activity at sub-millimolar range. The DPBA (**1**) is slightly more active than L-742,001 (**2**), as its efficacy appears from 250 µM whereas that of the L-742,001 (**2**) from 1 mM. Compounds DPBA (**1**) and L-742,001 (**2**) are both relatively fair binders but terribly weak inhibitors of the ENDO. However, both were described as potent inhibitors of Orthobunya- and Orthomyxo-endonuclease domain, at micromolar and sub-micromolar level. Their potency appeared relative to their mode of binding, which involves both ion chelation and hydrophobic or electrostatic interactions with some amino acids in the vicinity of the active site (Reguera *et al.*, 2010; Kowalinski *et al.*, 2012; DuBois *et al.*, 2012). The ENDO accommodates **1** and **2**, however the binding mode is only driven and predominately mediated by metal chelation. Indeed, if ions are removed, the binding is completely lost (Fig. 2). The lack of polar interactions between the compound and amino acids of the active site lead to residual flexibility and sub-optimal efficacy and makes them ligands but weak inhibitors. Our assumption in the use of DKAs, is that the compound compete with the RNA for the ions at the active site, thus creating a stable complex acting as a steric inhibitor. To design potent inhibitors, full occupancy binding mode should be obtained. In order to maximize our chances, we re-explored the Arenaviridae ENDO sequence structure conservation to guide the design of new DKAs able to establish hydrophobic interactions with amino acids of the ENDO active site.

As the chemical synthesis is more accessible, we chose the ligand **1** to guide the compound optimization and it will be the starting point for the structure-based drug-development.

3.6. Compound optimization and *in silico* evaluation

From a sequence point of view, *Arenaviridae* endonuclease has only 5 absolutely conserved regions from which 5 motifs can be derived. 4 of the 5 motifs (1 to 4) are located in

the vicinity of the catalytic site (Fig. 4). The central motif 2 with P-D-G, is the endonuclease catalytic signature. Motifs 1, 3 and 4 are surrounding motif 2 forming an hydrophobic slit below the catalytic site. Motif 1 is the furthest away but contributes together with the end of motif 2 by their hydrophobic nature to the formation of one side of the slit. Motifs 3 and 4 form the other side of the slit. They are both hydrophobic but punctuated with conserved polar amino acids including residues E101, K114, D118 and K121. These motifs show, despite years of evolution and viral diversity across the family, that the residues are invariant and consequently, highly important for the function or the fold integrity. Therefore for the compound optimization, these motifs can be used as solid ground to anchor compatible chemical moieties that would reinforce the binding and increase the inhibitory effect. In that spirit and after gaining insights into how DPBA (**1**) interacts with ENDO, we decided to explore the effect of modifications on the phenyl ring of DPBA (**1**) in an attempt to improve interactions with the amino acids of the active site and consequently binding efficacy.

We hypothesized that increasing the aromatic content of the ligands should reinforce the binding of the corresponding DKAs by establishing hydrophobic contacts with residues of motifs 3 and 4 (Fig. 4).

We designed two new DKAs, compounds **3** and **4**, which carry respectively, bi-phenyl, and phenantryl moieties (Fig. 1b). The selected aromatic rings reinforce the rigidity and hydrophobicity of the compound while increasing its electro-negativity, potentially allowing broader contacts in pockets. The bi-phenyl structure allows a larger degree of freedom compared to the phenantryl, allowing the second ring to rotate. Docking simulations confirmed our hypothesis, showing that increasing the number and type of aromatic components within the ligand (molecules **3** and **4**) a second "anchorage" into the targeted conserved area of the catalytic pocket (Fig. 5a, b and c). Poses of the different molecules have a binding energy (*E*) ranging between -7.2 and -5.5 Kcal/mol. This interaction can be described as follows: *i)* the DKA moiety interacts with the first metal ion through the carboxylic group and the enolizable α-ketone and the second metal ion through both α/γ-ketones; *ii)* the aromatic ring of compounds **3** and **4** generates hydrophobic interaction in conserved region motifs 3 and 4 with amino acids F103, V104, K114, Y146, and motifs 1, 3 and 4 with residues I49, E50, C102, F103, V104, R105, K114 respectively (Fig. 5d and S5).

In that perspective, we synthesized two new DKAs compounds **3** and **4.** These compounds were first evaluated for their affinity to ENDO-WT by MST, DKAs **3** and **4** display respectively K_d values of 0.05 ± 0.02 μM and 0.25 ± 0.07 μM. All two new DKAs show 10 to 100 times better affinities compared to DPBA (**1**), indicating that modifications afford additional interactions with ENDO-WT leading to better binding. We also tested the inhibition of DKAs **3** and **4** using our *in vitro* endonuclease activity assay (Fig. 5e and f). We observe that the intensity of the undigested RNA in presence of DKAs **3** and **4** remains

stronger compared to DPBA (**1**) (Fig. 5e). The quantification of the intensity of the bands reflects a diminution of the enzyme activity in presence of the new molecules **3** and **4**. This is an indication that compound modifications increasing hydrophobic interactions led to gain in potency (of about 35%, on the conservative side). These findings are the proof-of-concept that the design of optimized inhibitors, despite being challenging, is possible given that proper structural information is now available.

4. Conclusion

We report four crystallographic structures of the LCMV endonuclease domain: one in complex with two divalent ions (Mg^{2+}), the mutant D118A and the first two complexed structures with ligands : the DKAs DPBA (**1**) and L-742,001 (**2**). These structures provide the first detailed structural information of a specific ligand bound to an arenavirus endonuclease. Our data show that the presence of ions in holo-structure is in fact a rare event. Structures show that the metal ion binding depends on the number of water molecules able to create a mesh of interaction around the ions. This suggests that the catalytic ions responsible for the two-metal ion catalysis reaction are most likely brought in by the RNA substrate itself. This study shows that D88 is the critical residue that mediates metal ions binding. Moreover based on an extensive sequence analysis of all *Arenaviridae* endonuclease domains, we have identified 4 motifs that can be used as anchors for the specificity of optimized compounds targeting this domain. Our preliminary results, based on two new optimized DKAs, suggest that aromatic ring modifications improve binding and mostly anti-ENDO activity. In spite of the challenge of an open active site, this study is a proof of concept that the knowledge of ligand-bound structures can guide the design of optimized inhibitors against the LCMV endonuclease domain.

Figure 1 Structures of 2,4-dioxo-4-phenylbutanoic acid (DPBA) (**1**), (Z)-4-[1-benzyl-4-[(4-chlorophenyl)methyl]piperidin-4-yl]-2-hydroxy-4-oxobut-2-enoic acid (L-742,001) (**2**), 2-Hydroxy-4-(biphenyl-4-yl)-4-oxobut-2-enoic acid (**3**) and 2-Hydroxy-4-oxo-4-(phenanthren-3-yl)but-2-enoic acid (**4**).

Figure 2 (a) Thermal stability of ENDO proteins determined by Differential Scanning Fluorimetry (DSF). The melting temperature (T_m) of ENDO-WT, ENDO-D118A and ENDO-D88A (75 µM) (P), with or without the indicated divalent cation (0.5 mM) (P+Mg, P+Mn, P+Mg+Mn), and the compounds DPBA (**1**) (P+Mg+Mn+DPBA (**1**)) and L-742,001 (**2**) (P+Mg+Mn+L-742,001 (**2**)) (450 µM, ligand/protein ratio = 6) was measured in a Thermofluorescence experiment. (b) Polyacrylamide/8M urea gels of inhibition of endonuclease activity by DPBA (**1**) and L-742,001 (**2**). Increasing concentrations (µM) of molecules **1** and **2** were incubated with 20 µM of protein and 1 µM of single stranded RNA. The reaction products were analyzed in 20% polyacrylamide/8M urea gels. The NC

lane lacked protein in the reaction mixture meanwhile the EDTA lane had all reagents plus 5 mM EDTA.

Figure 3 Crystal structures of ENDO protein in complex with DPBA (**1**) and L-742,001 (**2**). LCMV endonuclease structure complexed with **1** (a-c) or **2** (d-f). Structures are represented as ribbon with helices in green and strands in gold while the compound is represented in sticks. Mg^{2+} and Mn^{2+} ions are represented as light green and purple spheres respectively. Panels (b) and (e) are a zoom of the catalytic site showing the intense water ions molecule catalytic residues coordination. Panels (c) and (f) show the 2Fo-Fc omit map corresponding to the compound (1σ).

Figure 4 The 5 motifs (blue) conserved across all the available sequences of *Arenaviridae* endonuclease domain plotted on LCMV structure (grey). Conserved motifs are represented as WebLogo. Size of the letter is representative of the frequency observed in the alignment (Probability of 1 being Identity). Numbering of the residue refers to the sequence of LCMV. Motifs 1 to 4 represents the 4 positions that can be used to anchor a ligand over all *Arenaviridae*. Star residues correspond to the key residues involved in active site or substrate binding and their side chain are highlighted on the structure (orange).

Figure 5 (a) Structure of ENDO-WT complexed with Mg ions and docking grid. Surface is coloured in cyan and conserved motif in orange. Mg^{2+} ions are represented as light green spheres. On the right panel is shown the position of DPBA (**1**). (b) Selection of the best docking poses of compounds **3** and **4**. (c) Zoom of best pose in the cavity surface (d) Zoom of best pose with the residues involved in binding. Those are coloured according to hydrophobicity from cyan (the least) to green (the most). (e) Polyacrylamide/8M urea gels of compounds **3** and **4** (50μM) incubated with 20 μM of protein and 1 μM of single stranded RNA. (f) Quantification of endonuclease activity in presence of DMSO and compounds **1**, **3** and **4**.

Table 1 Data-collection and refinement statistics for LCMV endonuclease domain complexed with ions, compounds, and mutant.

	LCMV ENDO complexed with Mg^{2+} ions	LCMV ENDO complexed with DPBA (**1**)	LCMV ENDO complexed with L-742,001 (**2**)	LCMV ENDO Mutant D118A
PDB code	5LTF	5LTN	5T2T	5LTS
Wavelength (Å)	0.976	0.978	0.873	0.978
Resolution range	48.46 - 2.43 (2.52 - 2.43)	53.59 - 1.88 (1.95 - 1.88)	54.06 - 1.97 (2.04 - 1.97)	53.57 - 2.51 (2.60 - 2.51)
Space group	P 41	P 41	P 41	C 2 2 21
Unit cell				
a,b,c	108.4, 108.4, 54.1	107.2, 107.2, 53.8	108.1,108.1, 54.1	144.8, 159.2, 52.7
α,β,γ	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Molecules in ASU	2	2	2	2
Total reflections	67793 (6444)	241187 (14092)	308124 (31957)	73597 (7143)
Unique reflections	23573 (2252)	49896 (4927)	44686 (4457)	21183 (2063)
Multiplicity	2.9 (2.9)	4.8 (2.9)	6.9 (7.2)	3.5 (3.4)
Completeness (%)	96.0 (96.0)	99.6 (99.5)	100 (100)	99.0 (100)
Mean I/sigma(I)	13.61 (1.58)	8.54 (1.20)	12.51 (2.02)	14.17 (2.09)
Wilson B-factor	57.12	37.53	33.03	58.27
R-merge	0.046 (0.607)	0.094 (0.868)	0.094 (0.870)	0.060 (0.594)
R-meas	0.056 (0.744)	0.105 (1.15)	0.102 (0.938)	0.071 (0.704)
CC1/2	0.999 (0.586)	0.996 (0.416)	0.998 (0.715)	0.997 (0.649)
No. of reflexions used in refinement (free reflections)	22999 (1932)	49844 (2457)	44682 (2210)	21091 (1004)
R-work	0.18 (0.27)	0.19 (0.29)	0.18 (0.27)	0.20 (0.27)
R-free	0.23 (0.33)	0.21 (0.29)	0.22 (0.30)	0.22 (0.29)
Total atoms in structure	3240	3353	3586	3184
macromolecules	3080	3064	3193	3063
Ligands	2	60	47	8
RMS (bonds)	0.009	0.010	0.008	0.013
RMS (angles)	1.09	1.10	1.86	1.67
Ramachandran favoured (%)	92	97	97	97
allowed (%)	5.7	1.9	2.3	2.7
outliers (%) *	2.3	1.1	0.7	0.3
Average B-factor	75.92	50.30	41.98	74.32
Beamline	ESRF ID23-1	Soleil Proxima I	ESRF ID23-2	Soleil Proxima I

Statistics for the highest-resolution shell are shown in parentheses.
*Outliers are in highly flexible regions of the protein.

Table 2 Determination of dissociation constants of ENDO proteins (WT, D118A and D88A) to DPBA (**1**) and L-742,001 (**2**) in the presence or absence of ions by Microscale Thermophoresis (MST). The concentration of labeled protein was kept constant at 100 nM, while the concentration of compound was varied from 500 μM to 15nM to yield 16 serial concentrations.

Compounds	K_D (μM)					
	WT		D118A		D88A	
	Ion	No ion	Ion	No ion	Ion	No ion
DPBA (1)	5.38 ± 2.06	ND	18.76 ± 3.88	ND	ND	ND
L-742,001 (2)	0.51 ± 0.11	ND	0.89 ± 0.06	ND	ND	ND

ND: not detectable.

Appendix A. Chemistry

All reagents and anhydrous solvents were purchased from Aldrich and used without further drying. All reactions involving air or moisture sensitivity compounds were performed under argon atmosphere, with oven-dried glassware and syringes. NMR spectra were determined in DMSO-*d6* on a JEOL spectrometer (400 MHz). Chemical shifts were expressed in ppm and coupling constants (*J*) are in hertz (s = singlet, d = doublet, m = multiplet). HPLC/MS spectra were recorded on an Accela 600 thermofisher (detection 260 nm) with an Hypersil-Gold C18 (2.5 µm, 2.1x50 mm) column and Finnigan survey MSQ (MS) under electrospray ionization (ESI) in positive and negative ionization mode detection. Compounds were eluted with a gradient of a solution of 5 to 100% of 0.01% formic acid in ACN solution with a 500 µL/min flow rate.

A.1. General procedure for the preparation of compounds 3 and 4.

Sodium (20 mmol) was added slowly to cold methanol in an ice/water bath to give a 2M solution of sodium methoxide. The appropriate aryl methyl ketone (5 mmol) and dimethyl oxalate (10 mmol)) were dissolved in dry THF/DME (1:1, v:v) and the solution was added dropwise to the freshly prepared sodium methoxide. The mixture was stirred at room temperature for 8 h. The resulting solid was filtered off, washed with cold methanol and diethyl ether and dried. The sodium ketoenolate ester obtained was dissolved in water and stirred at room temperature for 1h. Then, the solution was acidified to pH 3-4 by adding HCl 1N and kept at 4ºC for 2h. The precipitate was filtered off, washed with cold water and dried. The ketoenol ester obtained was dissolved in THF/MeOH (1:1, v:v) and a solution of NaOH 2N was added dropwise. The resulting clear solution was stirred at room temperature for 1h. The mixture was acidified to pH 2 by adding HCl 1N and the precipitate formed was filtered off, washed with cold water and dried under vacuum. The crude material obtained was purified by crystallizing it from diethyl ether which yielded the desired aryl diketo acids **3** and **4**.

A.1.1. 2-Hydroxy-4-(biphenyl-4-yl)-4-oxobut-2-enoic acid (3)

Obtained as a yellow solid (212 mg, 27% global yield) using 4-acetylbiphenyl as the starting material. Tr-HPLC/MS: 3.87 min. LC-MS (ESI^-): 266.97 $[M-H]^-$; HRMS ESI^- calcd for $C_{16}H_{12}O_4$ $[M-H]^-$ 267.0663, found 267.0663. ^1H NMR (400 MHz) (DMSO-*d6*) δ (ppm): 8.16 (d, 2H, *J* = 8.1 Hz), 7.88 (d, 2H, *J* = 8.1 Hz), 7.77 (d, 2H, *J* = 7.8 Hz), 7.54-7.50 (m, 2H), 7.47-7.38 (m, 1H), 7.15 (s, 1H), 4.60 (s, 2H, methylenic proton of the diketo tautomer, ketoenol/diketo ratio = 0.1). ^{13}C NMR (100 MHz) (DMSO-*d6*) δ (ppm): 189.73, 170.28, 163.16, 145.33, 138.61, 133.39, 129.14, 128.63, 127.24, 127.03, 97.87.

A.1.2. 2-Hydroxy-4-oxo-4-(phenanthren-3-yl)but-2-enoic acid (4)

Obtained as a yellow solid (185 mg, 23% global yield) using 3-acetylphenantrene as the starting material. Tr-HPLC/MS: 4.10 min. LC-MS (ESI^-): 290.94 $[M-H]^-$; HRMS ESI^- calcd for $C_{18}H_{12}O_4$ $[M-H]^-$ 291.0663, found 291.0664. ^1H NMR (400 MHz) (DMSO-*d6*) δ (ppm): 9.37 (s, 1H), 9.01 (s, 1H), 8.16-7.90 (m, 5H), 7.76-7.70 (m, 2H), 7.00 (s, 1H), 4.74 (s, 2H, methylenic proton of the diketo tautomer, ketoenol/diketo ratio = 0.15). ^{13}C NMR (100 MHz) (DMSO-*d6*) δ (ppm): 187.61, 174.57, 164.17, 134.01, 131.79, 131.79, 129.97, 129.37, 129.04, 128.72, 127.46, 127.37, 126.32, 124.89, 123.28, 122.52, 99.13.

Supporting Figures and Table legends.

Figure S1 Alignment of 245 sequences of *Arenaviridae* L ENDO domain prepared with ESPript.

Figure S2 *(a)* Representative binding curves for ENDO-WT to DPBA (**1**), L-742,001 (**2**), **3** and **4,** determined by microscale thermophoresis (MST). Fraction Bound represents the baseline corrected normalized fluorescence for each compound with normalization to 1 for amplitude. *(b)* Representative ITC titrations for ENDO proteins with DPBA (**1**) at 25 °C as a function of the molar ratio of ligand to protein. The upper panel shows the ITC raw data and the lower panel presents the integrated heats of each injection. Binding isotherms were fitted to the raw data using one-site binding model to determine the K_D.

Figure S3 (a) Zoom of the catalytic site of the holo-enzyme structure of the LCMV endonuclease with the 2 catalytic Mg^{2+} coordinated by D88 and a network of water molecules. (b) 2Fo-Fc Map of the D118A mutant. Overall structure of the endonuclease LCMV in not affected by this mutation.

Figure S4 Zoom of the catalytic site showing the intense water ions molecule catalytic residues coordination in stereo view. Crystal structures of ENDO protein in complex with (a) DPBA (**1**) and (b) L-742,001 (**2**). Structures are represented as ribbon with helices in green and strands in gold while the compound is represented in sticks. Mg^{2+} and Mn^{2+} ions are represented as light green and purple spheres respectively.

Figure S5 Zoom of best pose of molecules **3** (a) and **4** (b) in stereo view. Residues and their surface involved in binding are coloured according to hydrophobicity from cyan (the least) to green (the most). Structures are represented as ribbon while the compound is represented in sticks. Mg^{2+} ions are represented as light green spheres.

Table S1. Arenaviridae Endonuclease domain L NCBI reference number.

Funding - This work was supported by ANR grant ArenaBunya-L (ANR-11-BSV8-0019), the foundation pour la Recherche Médicale (FRM) (SPF20130526788), the Fondation Mediterranée Infection, and the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01).

Acknowledgements - The Authors would like to thanks the staff at synchrotrons SOLEIL beamline Proxima-1 and ESRF beamlines ID23-1 and ID23-2. Authors thanks Julie Lichière, Kévin Botelho Ferreira for their technical assistance, and Dr. Bruno Coutard, Dr. Juan Reguera and Dr. Dominga Rogolino for useful discussions.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367.

Barton, L. L., Mets, M. B. & Beauchamp, C. L. (2002). *Am. J. Obstet. Gynecol.* **187**, 1715–1716.

Bhatt, A., Gurukumar, K. R., Basu, A., Patel, M. R., Kaushik-Basu, N. & Talele, T. T. (2011). *Eur. J. Med. Chem.* **46**, 5138–5145.

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2210–2221.

Bonthius, D. J. (2012). *Semin. Pediatr. Neurol.* **19**, 89–95.

Bowen, M. D., Peters, C. J. & Nichol, S. T. (1997). *Mol. Phylogenet. Evol.* **8**, 301–316.

Briese, T., Paweska, J. T., McMullan, L. K., Hutchison, S. K., Street, C., Palacios, G., Khristova, M. L., Weyer, J., Swanepoel, R., Egholm, M., Nichol, S. T. & Lipkin, W. I. (2009). *PLoS Pathog.* **5**, e1000455.

Buchmeier, M. J., de la Torre, J.-C., Peters, C. J. & Torre, J. De (2007). *Fields Virology*, Vol. *II*, edited by D.M. Knipe & P.M. Howley, pp. 1791–1827. Philadelphia, PA, USA: Lippincott Williams & Wilkins.

Carcelli, M., Rogolino, D., Bacchi, A., Rispoli, G., Fisicaro, E., Compari, C., Sechi, M., Stevaert, A. & Naesens, L. (2014). *Mol. Pharm.* **11**, 304–316.

Charrel, R. N., Coutard, B., Baronti, C., Canard, B., Nougairede, A., Frangeul, A., Morin, B., Jamal, S., Schmidt, C. L., Hilgenfeld, R., Klempa, B. & de Lamballerie, X. (2011). *Antiviral Res.* **90**, 102–114.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). *J. Am. Chem. Soc.* **117**, 5179–5197.

Cornu, T. I. & de la Torre, J. C. (2001). *J. Virol.* **75**, 9415–9426.

Crooks, G., Hon, G., Chandonia, J. & Brenner, S. (2004). *Genome Res*. **14**, 1188–1190.

DuBois, R. M., Slavish, P. J., Baughman, B. M., Yun, M.-K., Bao, J., Webby, R. J., Webb, T. R. & White, S. W. (2012). *PLoS Pathog.* **8**, e1002830.

Edgar, R. C. (2004). *Nucleic Acids Res.* **32**, 1792–1797.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501.

Enria, D. A., Briggiler, A. M. & Sánchez, Z. (2008). *Antiviral Res.* **78**, 132–139.

Ferron, F., Weber, F., de la Torre, J. C. & Reguera, J. (2017). *Virus Res.* 1–17.

Fischer, S. A., Graham, M. B., Kuehnert, M. J., Kotton, C. N., Srinivasan, A., Marty, F. M., Comer, J. A., Guarner, J., Paddock, C. D., DeMeo, D. L., Shieh, W.-J., Erickson, B. R., Bandy, U., DeMaria, A., Davis, J. P., Delmonico, F. L., Pavlin, B., Likos, A., Vincent, M. J., Sealy, T. K., Goldsmith, C. S., Jernigan, D. B., Rollin, P. E., Packard, M. M., Patel, M., Rowland, C., Helfand, R. F., Nichol, S. T., Fishman, J. A., Ksiazek, T., Zaki, S. R. & Team, L. in T. R. I. (2006). *N. Engl. J. Med.* **354**, 2235–2249.

Fisher-Hoch, S. P., Tomori, O., Nasidi, a, Perez-Oronoz, G. I., Fakile, Y., Hutwagner, L. & McCormick, J. B. (1995). *BMJ*. **311**, 857–859.

Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.

Günther, S. & Lenz, O. (2004). *Crit. Rev. Clin. Lab. Sci.* **41**, 339–390.

Hastie, K. M., Igonet, S., Sullivan, B. M., Legrand, P., Zandonatti, M. A., Robinson, J. E., Garry, R. F., Rey, F. A., Oldstone, M. B. & Saphire, E. O. (2016). *Nat. Struct. Mol. Biol.* **23**, 513–521.

Hastie, K. M., Zandonatti, M., Liu, T., Li, S., Woods, V. L. & Saphire, E. O. (2016). *J. Virol.* **90**, 4556–4562.

Hinsen, K. (2000). *J. Comput. Chem.* **21**, 79–85.

Jamieson, D. J., Kourtis, A. P., Bell, M. & Rasmussen, S. A. (2006). *Am. J. Obstet. Gynecol.* **194**, 1532–1536.

Jiang, X., Huang, Q., Wang, W., Dong, H., Ly, H., Liang, Y. & Dong, C. (2013). *J. Biol. Chem.* **288**, 16949–16959.

Kilgore, P. E., Ksiazek, T. G., Rollin, P. E., Mills, J. N., Villagra, M. R., Montenegro, M. J., Costales, M. a, Paredes, L. C. & Peters, C. J. (1997). *Clin. Infect. Dis.* **24**, 718–722.

Kosata, B. & Danne, R. (2010). (http://bkchem.zirael.org/index.html . Version 0.13.0)

Kowalinski, E., Zubieta, C., Wolkerstorfer, A., Szolar, O. H. J., Ruigrok, R. W. H. & Cusack, S. (2012). *PLoS Pathog.* **8**, e1002831.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Crystallogr.* **26**, 283–291.

Lee, A. M., Pasquato, A. & Kunz, S. (2011). *Virology*. **411**, 163–169.

Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. (2010). *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958.

López, N., Jácamo, R. & Franze-Fernández, M. T. (2001). *J. Virol.* **75**, 12241–12251.

MacNeil, A., Ströher, U., Farnon, E., Campbell, S., Cannon, D., Paddock, C. D., Drew, C. P., Kuehnert, M., Knust, B., Gruenenfelder, R., Zaki, S. R., Rollin, P. E. & Nichol, S. T. (2012). *Emerg. Infect. Dis.* **18**, 1256–1262.

McCormick, J. B., King, I. J., Webb, P. A., Scribner, C. L., Craven, R. B., Johnson, K. M., Elliott, L. H. & Belmont-Williams, R. (1986). *N. Engl. J. Med.* **314**, 20–26.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Crystallogr.* **40**, 658–674.

Mets, M. B., Barton, L. L., Khan, A. S. & Ksiazek, T. G. (2000). *Am. J. Ophthalmol.* **130**, 209–215.

Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **65**, 1074–1080.

Morin, B., Coutard, B., Lelke, M., Ferron, F., Kerber, R., Jamal, S., Frangeul, A., Baronti, C., Charrel, R., De Lamballerie, X., Vonrhein, C., Lescar, J., Bricogne, G., Günther, S. & Canard, B. (2010). *PLoS Pathog.* **6**, e1001038.

Nicholson, A. W. (2014). *Wiley Interdiscip. Rev. RNA*. **5**, 31–48.

Noble, E., Cox, A., Deval, J. & Kim, B. (2012). *Virology*. **433**, 27–34.

Palermo, G., Cavalli, A., Klein, M. L., Alfonso-Prieto, M., Dal Peraro, M. & De Vivo, M. (2015). *Acc. Chem. Res.* **48**, 220–228.

Patil, S., Kamath, S., Sanchez, T., Neamati, N., Schinazi, R. F. & Buolamwini, J. K. (2007). *Bioorganic Med. Chem.* **15**, 1212–1228.

Perez, M., Craven, R. C. & de la Torre, J. C. (2003). *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12978–12983.

Perez, M., Greenwald, D. L. & de la Torre, J. C. (2004). *J. Virol.* **78**, 11443–11448.

Perez, M. & de la Torre, J. C. (2003). *J. Virol.* **77**, 1184–1194.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004*a*). *J. Comput. Chem.* **25**, 1605–1612.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004*b*). *J. Comput. Chem.* **25**, 1605–1612.

Pinschewer, D. D., Perez, M. & de la Torre, J. C. (2003). *J. Virol.* **77**, 3882–3887.

Reguera, J., Gerlach, P., Rosenthal, M., Gaudon, S., Coscia, F., Günther, S. & Cusack, S. (2016). *PLoS Pathog.* **12**, e1005636.

Reguera, J., Weber, F. & Cusack, S. (2010). *PLoS Pathog.* **6**,.

Rogolino, D., Carcelli, M., Sechi, M. & Neamati, N. (2012). *Coord. Chem. Rev.* **256**, 3063–3086.

Schulte, D. J., Comer, J. A., Erickson, B. R., Rollin, P. E., Nichol, S. T., Ksiazek, T. G. & Lehman, D. (2006). *Pediatr. Infect. Dis. J.* **25**, 560–562.

Steitz, T. a & Steitz, J. a (1993). *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6498–6502.

Stevaert, A., Dallocchio, R., Dessì, A., Pala, N., Rogolino, D., Sechi, M. & Naesens, L. (2013). *J. Virol.* **87**, 10524–10538.

Stevaert, A., Nurra, S., Pala, N., Carcelli, M., Rogolino, D., Shepard, C., Domaoal, R. A., Kim, B., Alfonso-Prieto, M., Marras, S. A. E., Sechi, M. & Naesens, L. (2015). *Mol. Pharmacol.* **87**, 323–337.

Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J. & Adams, P. D. (2008). *Acta Crystallogr. D. Biol. Crystallogr.* **64**, 61–69.

Trott, O. & Olson, A. J. (2010). *J. Comput. Chem.* **31**, 455–461.

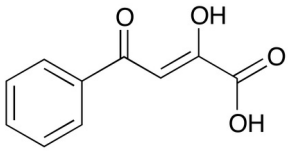
Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 293–302.

W.H.O (2016). Lassa Fever – Nigeria. (http://who.int/csr/don/27-may-2016-lassa-fever-nigeria/en/# . Accessed 08-08-2016)

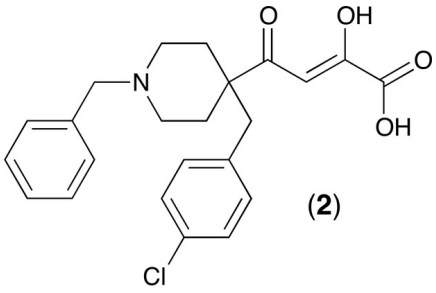
Wallat, G. D., Huang, Q., Wang, W., Dong, H., Ly, H., Liang, Y. & Dong, C. (2014). *PLoS One*. **9**, e87577.

Yun, N. E., Ronca, S., Tamura, A., Koma, T., Seregin, A. V., Dineley, K. T., Miller, M., Cook, R., Shimizu, N., Walker, A. G., Smith, J. N., Fair, J. N., Wauquier, N., Bockarie, B., Khan, S. H., Makishima, T. & Paessler, S. (2015). *J. Virol.* **90**, 2920–2927.

(a)

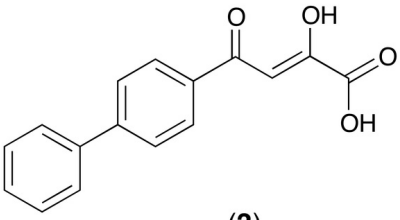


(**1**)

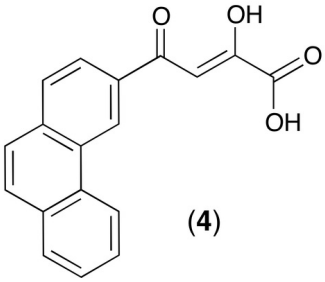


(**2**)

(b)



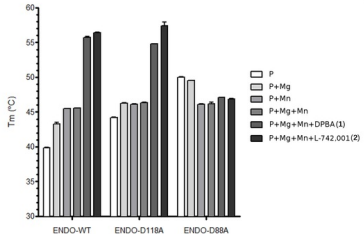
(**3**)



(**4**)

Figure 1

(a)


Tm (°C)
60
55
50
45
40
35
30
ENDO-WT
ENDO-D118A
ENDO-D88A
P
P+Mg
P+Mn
P+Mg+Mn
P+Mg+Mn+DPBA (1)
P+Mg+Mn+L-742,001 (2)

(b)

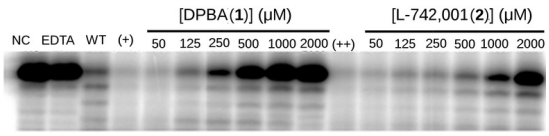

NC
EDTA
WT
(+)
[DPBA(1)] (μM)
50
125
250
500
1000
2000
(++)
[L-742,001(2)] (μM)
50
125
250
500
1000
2000

Figure 2

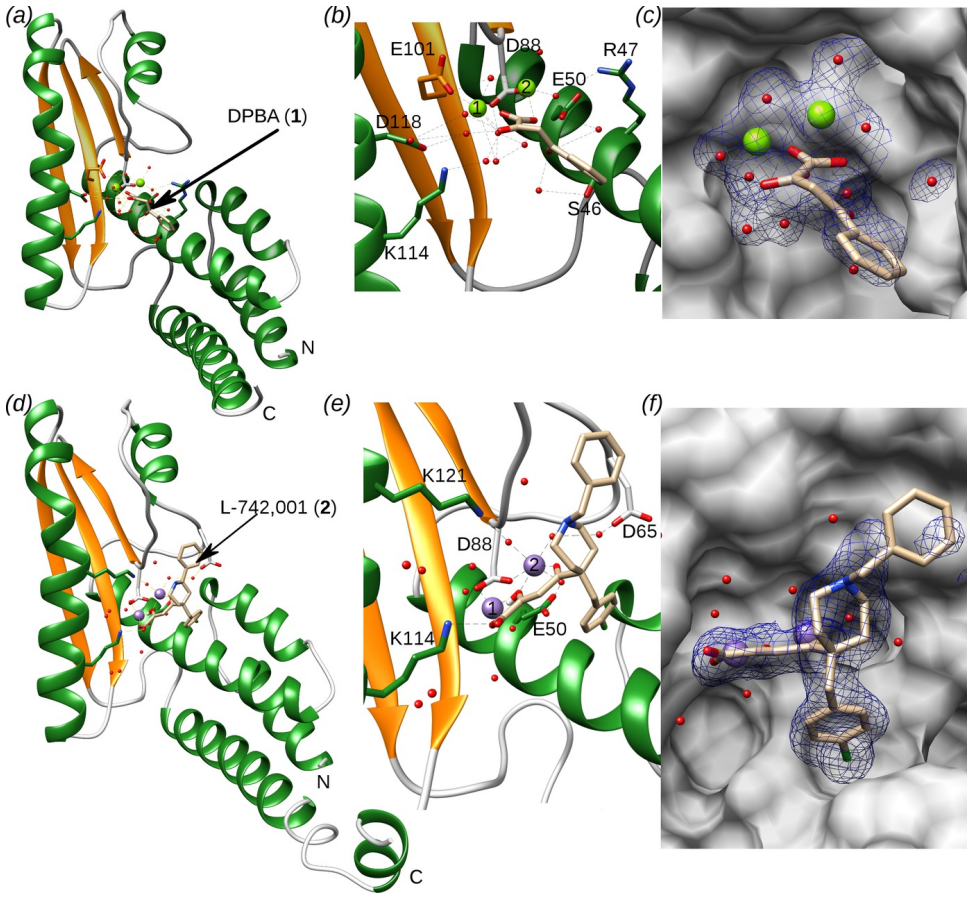

(a)
DPBA (1)
N
C
(b)
E101
D88
R47
E50
D118
S46
K114
(c)
(d)
L-742,001 (2)
N
C
(e)
K121
D88
D65
K114
E50
(f)

Figure 3

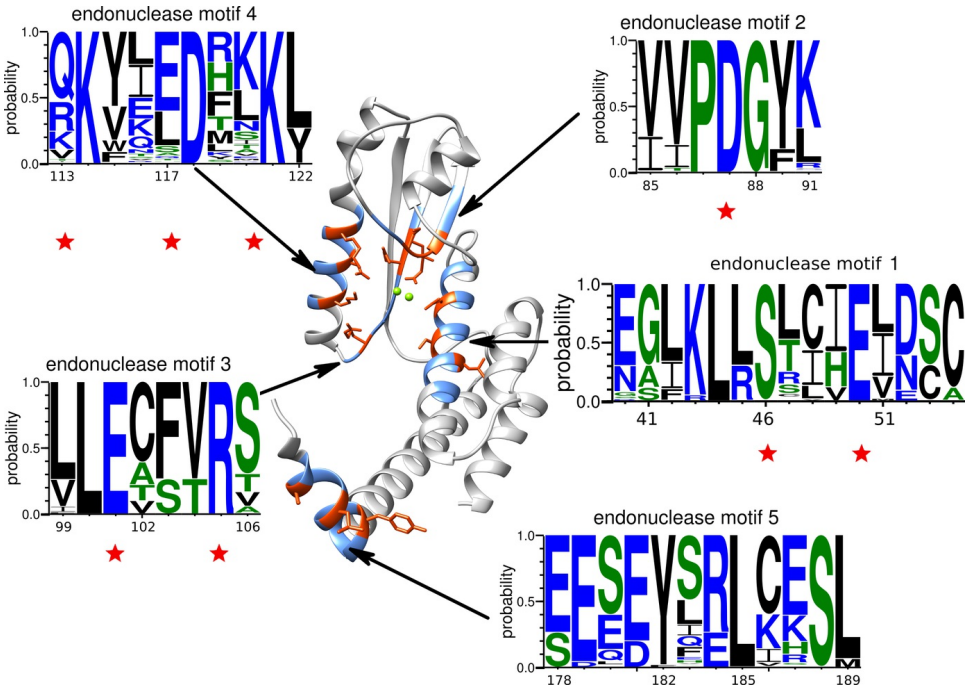

endonuclease motif 4
probability
1.0
0.5
0.0
113
117
122
endonuclease motif 2
85
88
91
endonuclease motif 1
41
46
51
endonuclease motif 3
99
102
106
endonuclease motif 5
178
182
185
189

Figure 4

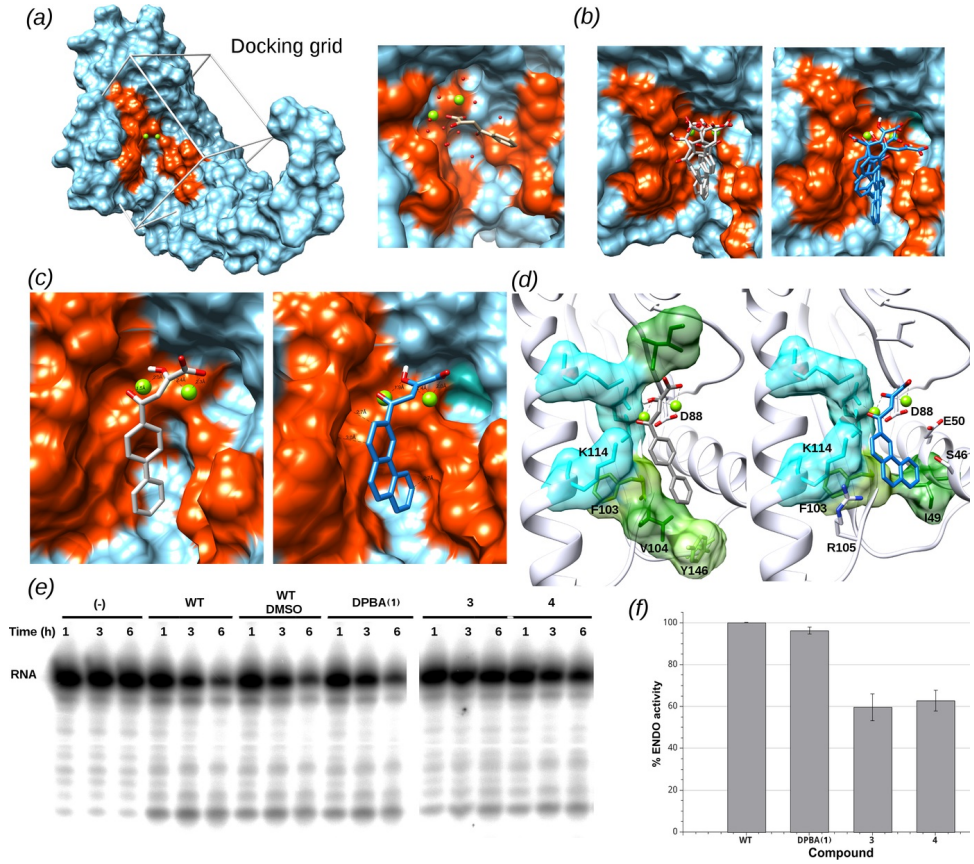

(a)
Docking grid
(b)
(c)
(d)
D88
K114
F103
V104
Y146
D88
E50
K114
S46
F103
I49
R105
(e)
(-)
WT
WT DMSO
DPBA(1)
3
4
Time (h) 1 3 6 1 3 6 1 3 6 1 3 6 1 3 6 1 3 6
RNA
(f)
% ENDO activity
100
80
60
40
20
0
WT
DPBA(1)
3
4
Compound

Figure 5