8-1-2020

Fragment screening targeting Ebola virus nucleoprotein C-terminal domain identifies lead candidates

David J Aceti

Hamza Ahmed

William M Westler

Chao Wu

Hesam Dashti

See next page for additional authors

Authors

David J Aceti, Hamza Ahmed, William M Westler, Chao Wu, Hesam Dashti, Marco Tonelli, Hamid Eghbalnia, Gaya K Amarasinghe, and John L Markley

Antiviral Research 180 (2020) 104822

Contents lists available at ScienceDirect

Antiviral Research

journal homepage: www.elsevier.com/locate/antiviral

Short Communication

Fragment screening targeting Ebola virus nucleoprotein C-terminal domain identifies lead candidates

David J. Aceti[a], Hamza Ahmed[b], William M. Westler[a], Chao Wu[b], Hesam Dashti[a,c], Marco Tonelli[a], Hamid Eghbalnia[a], Gaya K. Amarasinghe[b], John L. Markley[a,*]

[a] Biochemistry Department and National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, Madison, WI, 53706, USA
[b] Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, 63110, USA
[c] Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02215, USA

ARTICLE INFO

Keywords:
Ebola virus
Drug development
Nucleoprotein C-terminal domain
Fragment screening

ABSTRACT

The Ebola Virus is a causative agent of viral hemorrhagic fever outbreaks and a potential global health risk. The outbreak in West Africa (2013–2016) led to 11,000+ deaths and 30,000+ Ebola infected individuals. The current outbreak in the Democratic Republic of Congo (DRC) with 3000+ confirmed cases and 2000+ deaths attributed to Ebola virus infections provides a reminder that innovative countermeasures are still needed. Ebola virus encodes 7 open reading frames (ORFs). Of these, the nucleocapsid protein (eNP) encoded by the first ORF plays many significant roles, including a role in viral RNA synthesis. Here we describe efforts to target the C-terminal domain of eNP (eNP-CTD) that contains highly conserved residues 641–739 as a pan-Ebola antiviral target. Interactions of eNP-CTD with Ebola Viral Protein 30 (eVP30) and Viral Protein 40 (eVP40) have been shown to be crucial for viral RNA synthesis, virion formation, and virion transport. We used nuclear magnetic response (NMR)-based methods to screened the eNP-CTD against a fragment library. Perturbations of 1D ^{1}H NMR spectra identified of 48 of the 439 compounds screened as potential eNP CTD interactors. Subsequent analysis of these compounds to measure chemical shift perturbations in 2D ^{1}H,^{15}N NMR spectra of ^{15}N-labeled protein identified six with low millimolar affinities. All six perturbed an area consisting mainly of residues at or near the extreme C-terminus that we named "Site 1" while three other sites were perturbed by other compounds. Our findings here demonstrate the potential utility of eNP as a target, several fragment hits, and provide an experimental pipeline to validate viral-viral interactions as potential panfiloviral inhibitor targets.

1. Introduction

Ebola virus (EBOV) causes severe hemorrhagic fevers with case fatality rates up to 90% during rare outbreaks (Feldmann and Geisbert, 2011). The major 2013–2016 outbreak in West Africa (Messaoudi et al., 2015) and the more recent outbreak in the Democratic Republic of Congo (DRC) have motivated increased efforts to develop countermeasures. As yet, only a handful of therapies are under consideration for approval, with some recently receiving approval, while others are being used for emergency basis and under compassionate use guideline (Geisbert, 2017). Moreover, like other hemorrhagic fever viral infections, it appears that convalescent plasma and sera have shown some efficacy (Group et al., 2016). Recent efforts have mainly focused on therapeutics to target the Ebola Virus in three ways. The most popular and effective method is targeting the viral life cycle. Disruption of the cycle at key steps like RNA synthesis have been shown to be effective in in vitro and mice models, making small molecule inhibitors, antisense therapies like siRNA, and phosphorodiamate morpholino inhibitors as potent antiviral agents (Bixler et al., 2017). Other therapies have aimed to disrupt host-viral interactions, like the interaction of GP with host entry proteins like cathepsins (Bixler et al., 2017). Scientists have also looked to combat the disease via immune-modulatory therapies, like vaccines such as the Zmapp cocktail and therapeutics that can tame the cytokine storm and limit inflammation (Bixler et al., 2017). While the efforts on these different pipelines look promising, there is a clear gap in the small molecule approach, which can have significant advantages over vaccines and biologics, with their lower compliance burden and cold chain custody requirement. We describe here our efforts to target a

* Corresponding author. Biochemistry Department, 433 Babcock Dr., Univ. of Wisconsin-Madison, Madison, WI, 53706, USA.
E-mail addresses: djaceti@wisc.edu (D.J. Aceti), hahmed@wustl.edu (H. Ahmed), milo@nmrfam.wisc.edu (W.M. Westler), chaowu@wustl.edu (C. Wu), hdashti@bwh.harvard.edu (H. Dashti), tonelli@nmrfam.wisc.edu (M. Tonelli), eghbalni@nmrfam.wisc.edu (H. Eghbalnia), gamarasinghe@wustl.edu (G.K. Amarasinghe), jmarkley@wisc.edu, markley@nmrfam.wisc.edu (J.L. Markley).

https://doi.org/10.1016/j.antiviral.2020.104822
Received 10 January 2020; Received in revised form 8 May 2020; Accepted 15 May 2020
Available online 21 May 2020

key EBOV component by screening a small molecule fragment library.

The EBOV genome is a non-segmented, single-stranded, negative-sense RNA encoding seven open reading frames (ORFs) (Messaoudi et al., 2015). Among these ORFs, only the RNA-dependent RNA polymerase (RdRp) known as the Large or "L" protein, has enzymatic activity. While the L protein has so far resisted both structural determination and effective drug development, there are some broad-spectrum inhibitors. Therefore, an attractive alternative is the disruption of protein-protein or protein-RNA interactions. Drugs targeting interactions also have the purported advantage of slowing the development of resistance, since resistance may require a mutation disallowing drug binding and a second complementary mutation in the partner to restore the interaction (Cierpicki and Grembecka, 2015; Zinzalla and Thurston, 2009). Complete or partial structures of all six proteins have been solved, an enormous advantage in this effort. However, targeting interactions is challenging owing to the relatively flat, featureless surfaces often involved, as well as the relative novelty of this approach compared to targeting of enzyme active sites.

In this work we targeted the EBOV nucleoprotein (eNP), a multifunctional protein that is essential for genome replication, genome packaging, transcription, and viral structure. eNP binds the ssRNA viral genome and protects it from host nucleases, associates with proteins L/VP35/VP24/VP30 and genomic RNA to form the transcription/replication complex, and is essential for formation of the multiprotein nucleocapsid structure that surrounds the genome in the free virus (Messaoudi et al., 2015; Ruigrok et al., 2011).

The 739-residue eNP protein can be divided into structured hydrophobic N-terminal (1–450) and hydrophilic C-terminal domains (641–739) separated by a comparatively unstructured region (Dziubanska et al., 2014; Leung et al., 2015). The C-terminal domain (eNP-CTD) has no significant homology to any other protein; and its function is unclear; however, it is highly conserved among EBOV strains (Sherwood and Hayhurst, 2013); its deletion results in loss of viral viability (Muhlberger, 2007; Muhlberger et al., 1999); and antibodies cross-linked to it reduce viral progeny (Darling et al., 2017). It has been proposed that eNP-CTD is a hub for protein-protein interactions in the nucleocapsid (Dziubanska et al., 2014; Licata et al., 2004). Collectively, eNP-CTD and the unstructured region bind matrix protein VP40 (Licata et al., 2004; Noda et al., 2007) and transcription activator VP30 (Xu et al., 2017) and appear to regulate nucleocapsid formation, virion assembly, and budding (Bharat et al., 2012; Licata et al., 2004; Noda et al., 2007; Su et al., 2018).

High-resolution X-ray structures of eNP-CTD (PDB codes 4QB0 and 4QAZ) solved by Dziubanska and coworkers (Dziubanska et al., 2014) show a four α-helix/four β-strand arrangement distantly related to the β-grasp superfamily. It has been speculated that the extreme C-terminal helix is a critical central scaffold (Dziubanska et al., 2014).

Development of drug leads by screening with low molecular weight (< 200 Da) compound "fragments" instead of larger drug-like molecules is increasingly popular; much smaller chemical libraries are required for reasonable coverage of chemical space, and hit rates are much higher, typically 3–10% (Hubbard, 2016). However, such hits are weak binders (K_d high μM to low mM) and require extensive development, often including chemical linkage of fragments binding to neighboring sites, to develop drug-like affinity. Nuclear magnetic resonance (NMR) spectroscopy is one of only a few techniques that can detect such weak interactions.

We screened eNP-CTD against 439 fragment compounds. We used a one-dimensional ^{1}H NMR ligand-observe screen to quickly filter the chemical library to a manageable number of hits while requiring relatively small amounts of unlabeled protein. We carried out a secondary screen of the initial hits by two-dimensional ^{1}H,^{15}N Heteronuclear Single Quantum Coherence (HSQC) spectra of ^{15}N labeled protein in the absence and presence of added ligand, which determined binding location and affinity. Six chemically diverse compounds were found to bind to four sites on the protein surface with low mM affinity. All six compounds perturbed the chemical shifts of residues in the extreme C-terminus and those structurally nearby. Subsets of the six compounds perturbed residues located in three other sites. We chose from a separate fragment library 25 new compounds that were structurally similar to the best of the six hits, and screening of these identified several with K_d values of approximately 1 mM.

2. Materials and methods

2.1. NMR spectroscopy

All NMR spectra were acquired at 298 K (25 °C) on a Bruker Avance III 600 MHz spectrometer (Billerica, MA, USA) equipped with a ^{1}H (^{13}C/^{15}N/^{31}P) cryogenic probe SampleJet autosampler and Topspin v. 3.5p17 software. One-dimensional proton spectra were collected using the ZGESGPPE pulse sequence (Adams et al., 2013; Hwang and Shaka, 1995). MestreNova v12.0 software (Mestrelab Research S.L., Santiago de Compostela, Spain) was used to process and overlay spectra for comparison of peak intensities and frequencies; ^{1}H,^{15}N Heteronuclear Single Quantum Coherence (HSQC) perturbation and titration experiments were analyzed with the same software with the addition of the Chemical Shift Perturbation (Mbinding) plugin.

2.2. Preparation of fragment compound mixtures

The 1000-compound Maybridge Ro3 Core fragment library and the 2500-compound Maybridge Ro3 library were purchased from ThermoFisher Scientific (Waltham, MA, USA). We diluted 2.4 μL of each compound in plates 1–3 and 5–7 of the Core Library (483 total compounds supplied as 10 mM in 100% DMSO-d_6) to 160 μL with Screening Buffer A (11 mM bis-Tris-d19 buffer, pH 7.4, 150 mM NaCl, 0.04% NaN_3, with 15 μM DSS (Dimethyl-silapentane-sulfonate) and 200 μM Na sodium formate as chemical shift references, in D_2O) to achieve a final concentration of 150 μM. For collection of ^{1}H NMR reference spectra, solutions were transferred into 3 mm Bruker SampleJet tubes. We developed MestreNova scripts for high throughput spectral processing and to generate peak lists used as input to NMRMix (Stark et al., 2016) for the virtual assembly of mixtures with minimal peak overlap (93 mixtures of 4–5 compounds each). We used another in-house script (H. Dashti, unpublished) that converted the output from NMRMix to a worklist for an Eppendorf (Hauppauge, NY) epMotion 5075 liquid-handling robot, which physically assembled mixtures of the compounds so that each was 2 mM in DMSO-d_6.

2.3. Protein expression and purification

Both unlabeled and ^{15}N labeled eNP (641–739) proteins were expressed as maltose binding protein (MBP) fusion proteins in BL21(DE3) *E. coli* cells (Novagen). For unlabeled protein, LB media was used. For ^{15}N labeling, M9 media was used supplemented with 1.5 g/L ^{15}N NH_4Cl, 2 g/L glucose, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, and 1X vitamin mix (Sigma-Aldrich B6891). Protein expression was induced at an optical density of 0.6 (measured at 600 nm) with 0.5 mM IPTG and grown for 12–15 h at 18 °C. Cells were harvested, resuspended in lysis buffer containing 25 mM Tris (pH 7.5), 150 mM NaCl, 20 mM imidazole, and 5 mM 2-mercaptoethanol (BME), lysed using an EmulsiFlex-C5 homogenizer (Avestin), and clarified by centrifugation at 30,000 × *g* at 4 °C for 45 min eNP 641–739 constructs were purified using a series of affinity and ion exchange chromatographic columns. The MBP tag was cleaved using TEV protease prior to a final application on a size exclusion column. The purity of the samples was determined by SDS-PAGE.

2.4. Initial ligand-observe fragment screening by ^{1}H T2 relaxation NMR

Protein was buffer-exchanged by centrifugal filtration into

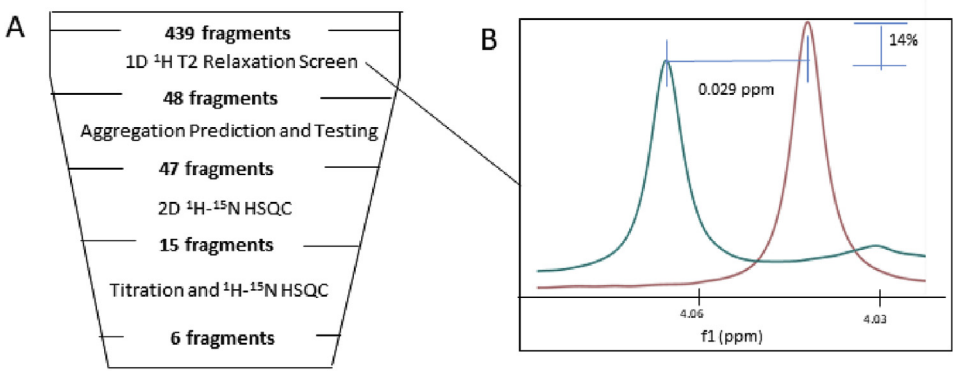


Fig. 1. Identification of fragments binding to eNP-CTD. (A) The C-terminal domain of eNP (residues 641–739) was screened with 439 fragment compounds from the Maybridge Ro3 Core Library, formulated as mixtures of 4–5 compounds such that 15 μM protein was exposed to 75 μM of each compound. **(B)** Example of overlaid one-dimensional proton spectra of a compound peak with 15 μM protein (blue) and without (red). Significant protein-binding effects, defined as ≥ 10% loss of peak intensity or ≥0.002 ppm (1.2 Hz) change in chemical shift or both, for at least one compound peak, were found for 48 compounds.

Screening Buffer A (described above) or Screening Buffer B (50 mM phosphate buffer, pH 7.4, 75 mM NaCl, 0.04% NaN_3, 15 μM DSS, in D_2O); the protein concentration in these buffers was adjusted to 15 μM. Fragment compound mixtures were diluted from 6 μL to a final 160 μL with screening buffer (for samples without protein) or with 15 μM protein in the same buffer (for samples with protein). Final concentrations were 75 μM for each compound and 3.75% v/v for DMSO-d_6.

2.5. Secondary protein-observe screening by $^1H,^{15}N$ HSQC NMR

We utilized 2D $^1H,^{15}N$ HSQC NMR spectra of ^{15}N-labeled protein to detect chemical shift perturbations from backbone and side-chain amide groups resulting from the addition of hit compounds from the initial screen. Experiments were performed in 10 mM HEPES buffer, pH 7.0, 150 mM NaCl, 0.04% NaN_3, with 10% D_2O for the frequency lock. Perturbation tests were performed with compounds at 1628 μM and protein at 50–100 μM. Titrations to determine affinity were performed with compounds at 25, 51, 153, 407, 814, and 1628 μM, and protein at 35–100 μM. Significant perturbation for individual residues was defined as ≥ 0.015 ppm combined $^{15}N/^1H$ perturbations, using MestreNova's default normalization factor (Eq. (1)):

$$([\Delta\delta_H]^2 + [0.156*\Delta\delta_N]^2)^{0.5} \quad (1)$$

We combined chemical shift perturbation data from multiple titration experiments on residues within discrete binding sites to determine K_d values and standard errors for each compound and binding site. We used SigmaPlot 13.0 (Systat Software, Inc., San Jose, CA) to analyze the data according to "Simple Ligand Binding" and "one site saturation" (Eq (2)):

$$y = B_{max}x/K_d + x \quad (2)$$

2.6. FTMap analysis

We used the FTMap computational map server (Kozakov et al., 2015) to identify ligand-binding hotspots. The NP-CTD structure 4QB0 (Dziubanska et al., 2014) was uploaded, and analysis run in PPI (Protein-Protein Interaction) Mode, according to instructions. The results were visually inspected using PyMol (DeLano, 2002).

2.7. Virtual ligand docking

We used the docking program HADDOCK (van Zundert et al., 2016) according to instructions for "Small Molecule Binding Site Screening" to predict ligand binding locations and poses. eNP-CTD structure 4QBO was manually formatted to HADDOCK specifications. Perturbed eNP-CTD residues identified in this work were used in an initial step as unambiguous restraints to place the ligand within binding sites, then as ambiguous restraints to allow the molecule to move within the site.

3. Results and discussion

3.1. Testing fragment compounds and assembling mixtures

We purchased a compound fragment screening library to test for binding against recombinant NP. To validate the identity and purity of the compounds from Maybridge, we used the provided structure along with the 1D ^1H NMR spectrum of each of 483 compounds as input to GISSMO software (Dashti et al., 2017b, 2018bib_Dashti_et_al_2018bib_Dashti_et_al_2017b) used in simulating its spin system matrix. We used in-house software [H. Dashti, an unpublished extension to ALATIS (Dashti et al., 2017a)] to overlay the spectrum simulated from the spin system matrix with the experimental spectrum. This analysis identified 439 of the compounds as chemically correct and free of major contaminants (all spectra and their corresponding spin system matrices are available at http://gissmo.nmrfam.wisc.edu/library by chemical name and by Maybridge Ro3 Core Library plate positions). Confirmed compounds were assembled into mixtures as described in Materials and Methods.

3.2. Initial screening of eNP-CTD protein against fragment compounds

We collected 1D ^1H NMR spectra of the fragment compound mixtures without and with unlabeled eNP-CTD protein. Comparison of overlaid spectra indicated 48 compounds exhibiting binding-induced spectral perturbations (Fig. 1).

One of the 48 compounds, Maybridge CC03846, was predicted by online tools ZINC Patterns Identifier (Irwin et al., 2015) and Dock Advisor (Ferreira et al., 2010) to self-aggregate, which often results in non-specific binding and false positive results (Dahlin et al., 2015; Ferreira et al., 2010)). Experimental confirmation was obtained as disaggregation-induced changes in chemical shifts, peak dispersion, and peak intensity in 1D ^1H NMR spectra as a function of progressive dilution or addition of 0.16% Triton X-100 detergent (LaPlante et al., 2013) (data not shown) and the compound was discarded.

3.3. Secondary screen employing ^{15}N-labeled eNP-CTD protein

In order to determine which residues these remaining 47 compounds perturbed, the remaining compounds were tested individually against ^{15}N-labeled protein at either a single 1628 μM compound concentration *vs.* a DMSO-d_6 control or with a titration series between 25 μM and 1628 μM (Fig. 2) in which DMSO-d_6 was held constant at 0.8%. $^1H,^{15}N$ HSQC protein spectra showed 20 compounds (not shown) for which at least one residue was perturbed. Five compounds perturbing only a tryptophan sidechain amide were discarded as making insufficient contact with the protein.

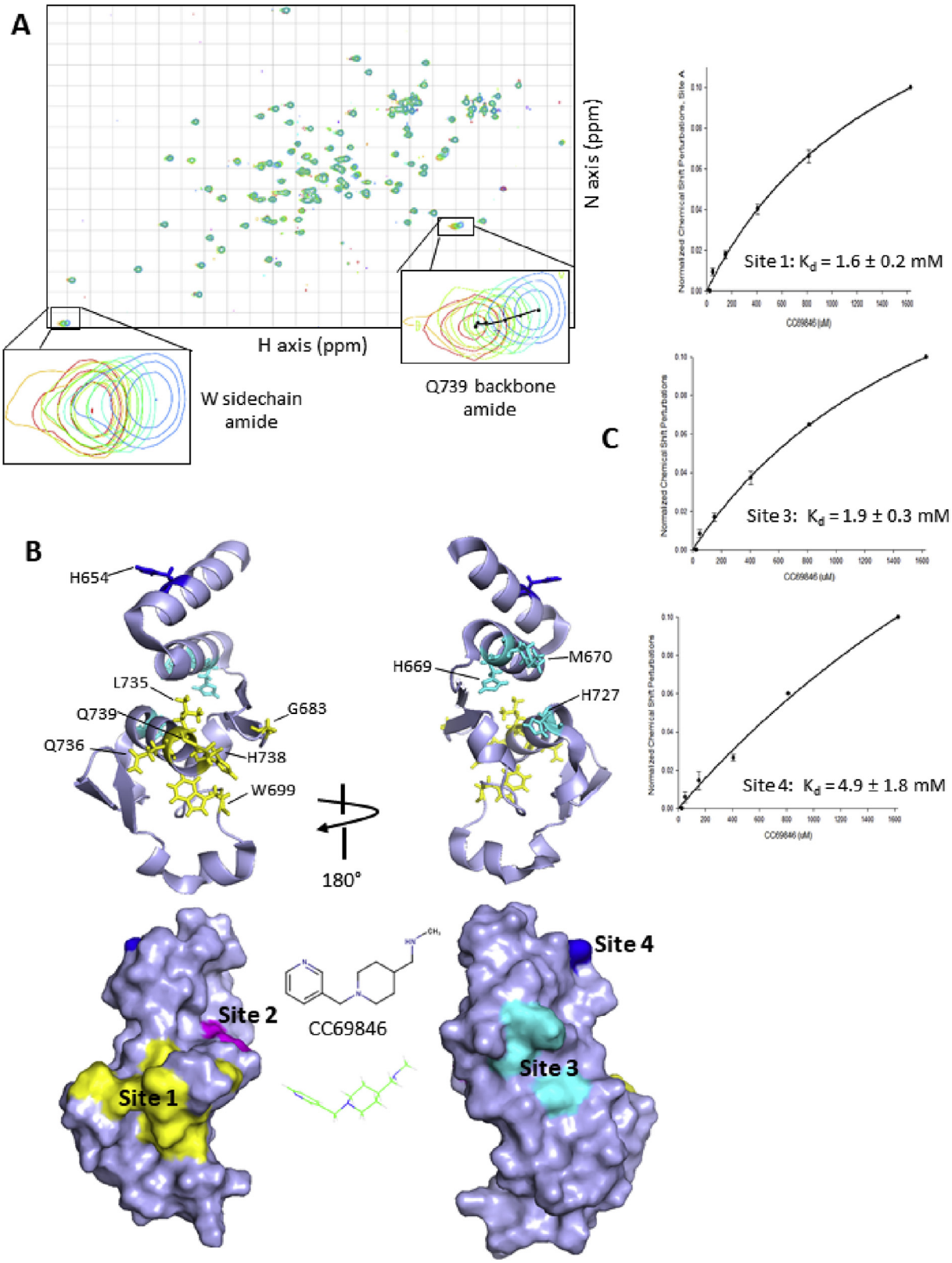


Fig. 2. Affinity determination by ^{1}H,^{15}N HSQC NMR and mapping of perturbations onto protein structure. (A) Forty-seven compounds were tested by ^{1}H,^{15}N protein-observe HSQC experiments (shown here, Maybridge CC69846) at single concentrations or in titration series and significantly perturbed peaks noted. The Q739 backbone amide and an unassigned W (Trp) sidechain amide are shown as examples of commonly perturbed peaks. **(B)** The nine residues perturbed by CC69846 could be grouped into Sites 1, 3, and 4; Site 2 was perturbed by other compounds. Structure models were prepared using Pymol. **(C)** Titration data for all perturbed residues within a site for three replicate CC69846 titrations were combined and fitted to give a single binding curve and K_d for each site.

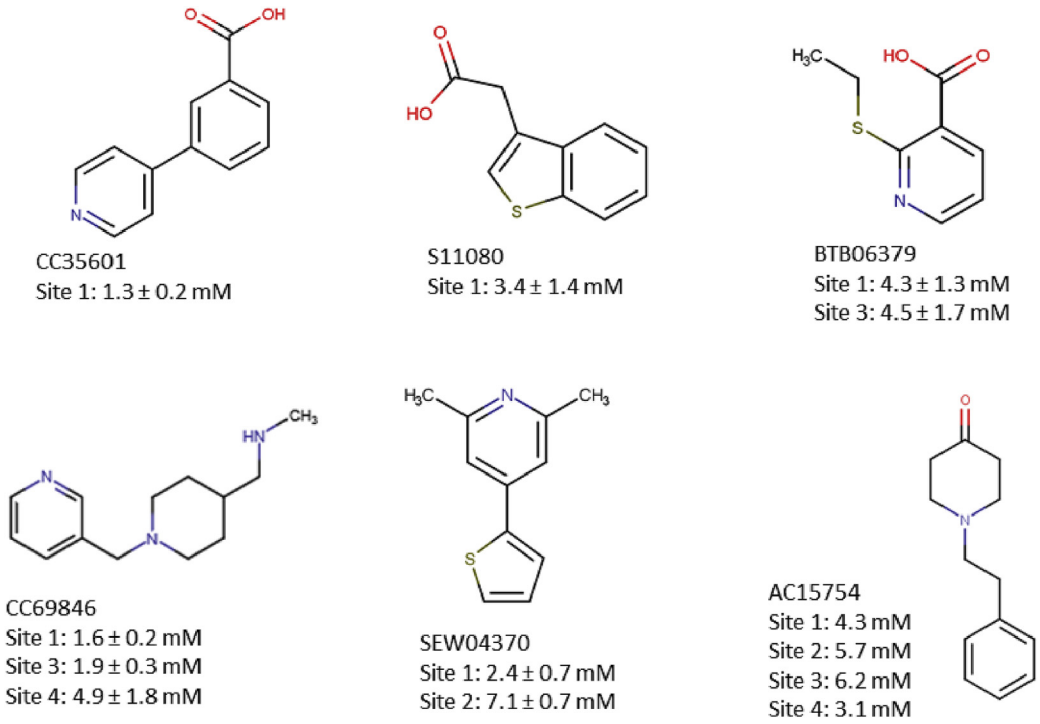


Fig. 3. Six compounds binding to discrete sites with K_d < 7.0 mM. Binding sites and K_d values for each compound are shown.

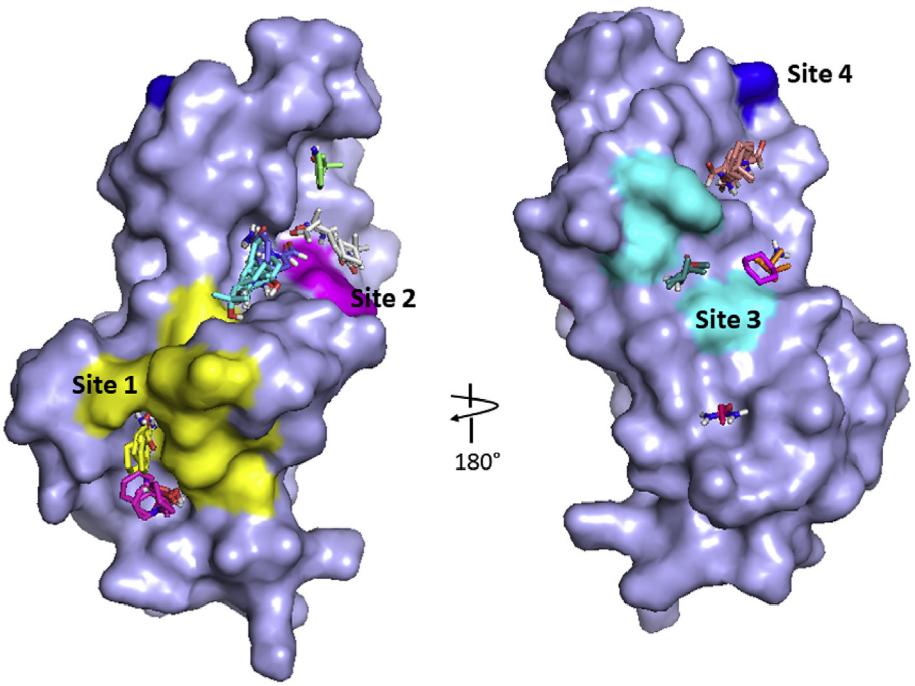
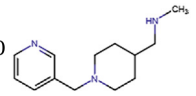
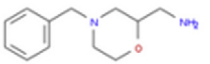
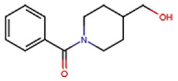
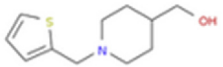
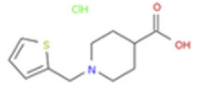
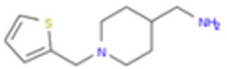
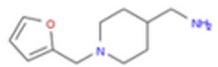
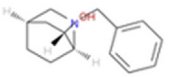
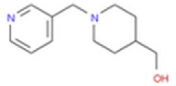
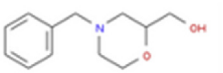
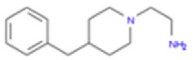
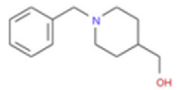
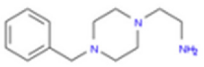
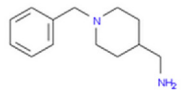
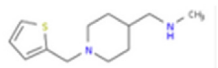
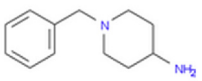
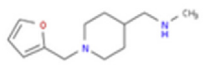
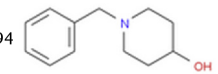


Fig. 4. Correspondence of small-molecule binding hotspots with perturbed sites. FTMap placed up to 16 compounds onto various sites on the crystal structure of eNP-CTD (4QB0) (Dziubanska et al., 2014) to identify small-molecule binding hotspots. For comparison, commonly perturbed residues at the four experimentally-determined binding sites are indicated by colored patches.

3.4. Mapping perturbations onto the eNP-CTD protein structure and affinity measurements

Residues perturbed by the remaining 15 compounds were mapped onto X-ray structure 4QBO (Dziubanska et al., 2014) using the ^{1}H,^{15}N HSQC NMR assignments from the same source (Fig. 2). Binding affinities were determined by titration of ^{15}N-labeled protein with each compound up to 1628 μM. Perturbations mapping onto discrete sites were plotted as a group *vs.* compound concentration to determine a single affinity for that site. Only six compounds (Fig. 3) had affinities with K_d < 7.0 mM as determined from chemical shift perturbations at one or more of four discrete patches on the surface of the protein. All six compounds perturbed Site 1, consisting of extreme C- terminal residues commonly including W699, L735, Q736, H738, and Q739, as well as structurally adjacent residues (G683, in the example shown in Fig. 2, or T714). Two compounds additionally perturbed Site 2 (shown in Fig. 2B), most commonly consisting of residues V676 and V677. Three compounds perturbed Site 3, most commonly residues H669, M670, and H727. Two compounds perturbed Site 4, in which only H654 was significantly affected.

Table 1
Binding locations, affinities, and ligand efficiencies for NP-CTD of CC69846 and related compounds.

Compound	Site 1[a]	Site 2	Site 3	Site 4	Compound	Site 1	Site 2	Site 3	Site 4
CC69846 (seed) 	**1.6 ± 0.2 (0.24)**	**None**[b]	**1.9 ± 0.3 (0.24)**	**4.9 ± 1.8 (0.20)**	CC12413 	2.1 ± 0.2 (0.21)	2.9 ± 1.2 (0.22)	2.8 ± 0.6 (0.22)	2.6 ± 0.9 (0.23)
CC70209 	None	None	None	None	CC66709 	3.2 ± 0.8 (0.22)	None	None	None
CC66751 	None	None	None	None	CC66713 	2.0 ± 0.4 (0.24)	None	1.5 ± 0.4 (0.25)	1.9 ± 0.6 (0.24)
CC68513 	1.7 ± 0.3 (0.24)	None	1.1 ± 0.2 (0.26)	2.3 ± 0.7 (0.23)	BTBG00045 	2.4 ± 0.3 (0.23)	3.8 ± 2.8 (0.21)	None	5.0 ± 3.6 (0.20)
CC69809 	2.1 ± 0.4 (0.23)	None	None	3.0 ± 1.3 (0.22)	CC12409 	2.9 ± 2.5 (0.22)	None	None	None
MO00719 	1.3 ± 0.2 (0.25)	4.0 ± 1.0 (0.21)	1.6 ± 0.3 (0.24)	2.2 ± 0.5 (0.23)	KM09623 	2.7 ± 2.6 (0.22)	None	None	None
KM10049 	2.4 ± 0.6 (0.23)	None	2.4 ± 0.6 (0.21)	3.7 ± 1.9 (0.21)	MO00973 	2.2 ± 0.6 (0.23)	None	3.1 ± 1.0 (0.22)	3.6 ± 1.3 (0.21)
CC66746 	1.5 ± 0.3 (0.25)	None	1.7 ± 0.4 (0.24)	1.6 ± 0.3 (0.24)	SB01177 	2.1 ± 0.4 (0.23)	4.0 ± 3.0 (0.21)	2.0 ± 0.5 (0.24)	1.8 ± 0.3 (0.24)
CC68546 	1.5 ± 0.2 (0.25)	None	1.6 ± 0.2 (0.24)	2.1 ± 0.4 (0.23)	JFD00694 	4.0 ± 1.4 (0.21)	None	5.8 ± 2.3 (0.20)	6.9 ± 2.9 (0.19)

[a] K_d, mM ± Std. Error (ligand efficiency).
[b] None = No residues in that site were perturbed, or titration resulted in Kd ≥ 7.0 mM.

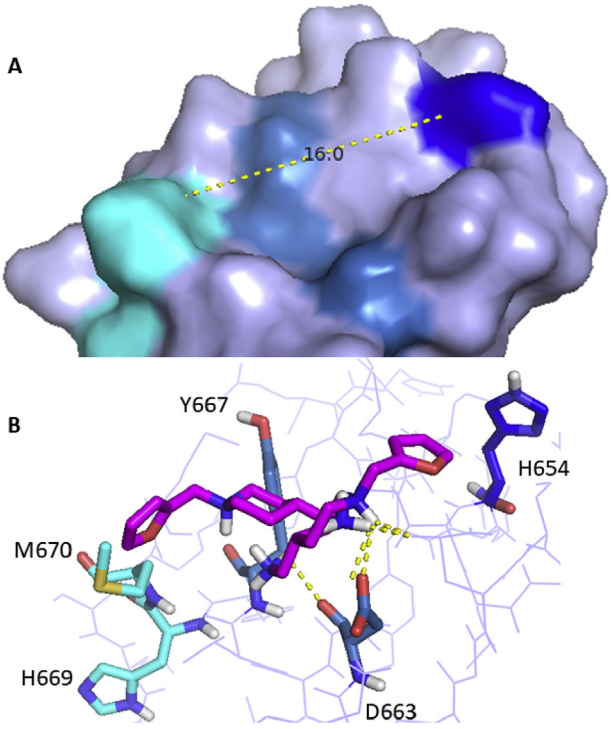


Fig. 5. Groove between Sites 3 and 4 and Docking of CC68513. (A) The shallow groove between Site 3 (cyan, H669 and M670) and Site 4 (dark blue, H654) with moderately perturbed residues D663 and Y667 (medium blue) is shown with the distance between M670 and H654 side chains. **(B)** Two poses of compound CC68513 (magenta) placed within the groove by docking program HADDOCK are shown in the same orientation as above, with hydrogen bonds to D663 and Q659 (the latter not perturbed or highlighted) as predicted by Pymol. Compound CC68513 itself is 9.5 Å long. Three other poses suggested by HADDOCK have been omitted for clarity.

3.5. Comparison of perturbation sites to predicted small-molecule binding hotspots

FTMap is a tool for identifying small-molecule binding hotspots (Kozakov et al., 2015) on the basis of fitting small compounds of varying size, shape and polarity onto the surface of a protein structure model as scored by calculated van der Waals and electrostatic energies as well as shape compatibility. FTMap assigned the highest hotspot scores for eNP-CTD protein (PDB: 4QBO) to the groove between Sites 1 and 2 (Fig. 4, cyan stick representations) and at the lower of end of Site 1 (pink sticks) where it placed all 16 compounds. FTMap assigned a high score to the deep pocket in Site 1 (yellow sticks) that was perturbed by 15 of the compounds.

3.6. Testing of compounds similar to CC35601 (3-pyrid-4-ylbenzoic acid)

We used ChemMine (http://chemmine.ucr.edu/) Hierarchical Clustering of Tanimoto scores to identify eight compounds from a new library, the 2500-compound "Maybridge Ro3", related to CC35601 with the goal of increasing binding affinity. We determined the effects of these on $^{1}H,^{15}N$ HSQC spectra of ^{15}N-labeled eNP-CTD at single concentrations (1628 μM). All eight compounds were found to bind on the basis of peak perturbations. However, titration studies showed that none had stronger affinity than compound CC35601 (Suppl. Table 1).

3.7. Testing of compounds similar to CC69846 (N-methyl-[1-(pyrid-3-ylmethyl)piperid-4-yl]methylamine)

CC69846 perturbed Sites 1 and 3 with K_d 1.6 ± 0.2 mM and K_d 1.9 ± 0.3 mM, respectively, as well as Site 4 with lower affinity. Motivated by the possibility of improved affinity at any of the three sites, we used ChemMine to select the 17 compounds in the 2500-compound library most resembling CC69846 (Table 1). Titration effects of these compounds on $^{1}H,^{15}N$ HSQC spectra of ^{15}N-labeled eNP-CTD showed that positively charged amines on the right ring sidechain appear to increase affinity while negatively charged carboxylates abolish binding and neutral hydroxyls decrease affinity somewhat. We found no compounds that bound more tightly than CC69846 at Site 1. Although seed compound CC69846 did not bind Site 2, four of the related compounds did so.

Thirteen of the 16 compounds that bound NP-CTD did so at multiple sites, a possible indicator of nonspecific binding. However, 2 compounds with high similarity to CC69846 did not bind at all to the protein (Table 1). In addition, only 2 of the 16 compounds in Table 1 (CC68513 and CC66709) bound an unrelated Ebola protein (eVP30, residues 130–272) in single-compound ligand-observe NMR experiments like those in Fig. 1 (data not shown). Both of these factors argue against promiscuity.

Affinities at Sites 3 and 4 exhibited some correlation, suggesting that they may be considered a single site. Indeed, residue M670 in Site 3 and residue H654 in Site 4 are only 16 Å apart (Fig. 5). FTMap identification of a single hotspot containing 14 of 16 possible compounds between these sites (Fig. 4) tends to support this hypothesis. Perturbation-guided docking of compound CC68513 using Haddock (Geng et al., 2017) suggests two high-quality binding poses with predicted hydrogen bonding in this pocket (Fig. 5).

4. Discussion

Previous studies have highlighted the importance of the eNP CTD in context of viral replication and potentially impacting protein interactions, in part through truncation analysis (Bharat et al., 2012; Licata et al., 2004). But, we currently lack direct means to target eNP CTD and its interactions critical for viral replication through small molecules. To address this need, we screened a fragment library from which we identified several fragment hits that potentially interact with specific regions with in the eNP CTD. Interestingly, eNP CTD has been implicated in interactions with eVP40 and other minor matrix proteins such as eVP30 and eVP24 (Watanabe et al., 2006). Additionally, several studies have labeled the eNP CTD as a hub for interaction with eVP40, the matrix protein important for proper assembly and transport of the nucleocapsid. These interactions were determined through a co-immunoprecipitation experiment between both WT eNP FL and WT eVP40, but when the CTD was truncated by 50 residues, the binding was abolished. The truncation analysis points to the possibility of the binding domain residing within the eNP CTD screened against above, but a structural change in residues closer to the N-terminus within the CTD (not contained in the 641–739 region) could similarly explain the drop in binding (Watanabe et al., 2006). Validating this interaction could prove significant given that the eVP40-eNP interactions have been shown to be necessary for efficient transport and nucleocapsid incorporation into viral capsids. Our results now provide an opportunity to test these important interactions mediated by eNP CTD.

Our studies here, using NMR-based fragment screening assays have identified several fragments that can potentially bind and restrict interactions between eNP CTD with viral and host proteins. Our data also identifies several regions in the eNP CTD that may be important for forming the viral nuclearcapsid and for viral replication. In addition, our fragment-based screens also provide the framework for future studies to tether the compounds in order to target multiple sites within the eNP CTD.

Declaration of competing interest

None.

Acknowledgements

Supported by NIH R01AI123926 (JLM and GKA) and P01AI120943 (GKA). NMR studies were carried out at the National Magnetic Resonance Facility at Madison (NMRFAM) which is supported by NIH grant P41GM103399. H.D. is supported in part by National Heart Lung and Blood Institute T32 HL007575.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.antiviral.2020.104822.

References

Adams, R.W., Holroyd, C.M., Aguilar, J.A., Nilsson, M., Morris, G.A., 2013. Perfecting" WATERGATE: clean proton NMR spectra from aqueous solution. Chem. Commun. 49, 358–360.

Bharat, T.A., Noda, T., Riches, J.D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoka, Y., Briggs, J.A., 2012. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. Proc. Natl. Acad. Sci. U. S. A. 109, 4275–4280.

Bixler, S.L., Duplantier, A.J., Bavari, S., 2017. Discovering drugs for the treatment of ebola virus. Curr. Treat. Options Infect. Dis. 9, 299–317 2017.

Cierpicki, T., Grembecka, J., 2015. Targeting protein-protein interactions in hematologic malignancies: still a challenge or a great opportunity for future therapies? Immunol. Rev. 263, 279–301.

Dahlin, J.L., Nissink, J.W., Francis, S., Strasser, J.M., John, K., Zhang, Z., Walters, M.A., 2015. Post-HTS case report and structural alert: promiscuous 4-aroyl-1,5-disubstituted-3-hydroxy-2H-pyrrol-2-one actives verified by ALARM NMR. Bioorg. Med. Chem. Lett 25, 4740–4752.

Darling, T.L., Sherwood, L.J., Hayhurst, A., 2017. Intracellular crosslinking of filoviral nucleoproteins with xintrabodies restricts viral packaging. Front. Immunol. 8, 1197.

Dashti, H., Wedell, J.R., Westler, W.M., Tonelli, M., Aceti, D., Amarasinghe, G.K., Markley, J.L., Eghbalnia, H.R., 2018. Applications of parametrized NMR spin systems of small molecules. Anal. Chem. 90, 10646–10649.

Dashti, H., Westler, W.M., Markley, J.L., Eghbalnia, H.R., 2017a. Unique identifiers for small molecules enable rigorous labeling of their atoms. Sci Data 4, 170073.

Dashti, H., Westler, W.M., Tonelli, M., Wedell, J.R., Markley, J.L., Eghbalnia, H.R., 2017b. Spin system modeling of nuclear magnetic resonance spectra for applications in metabolomics and small molecule screening. Anal. Chem. 89, 12201–12208.

DeLano, W.L., 2002. The PYMOL Molecular Graphic System. DeLano Scientific LLC, San Carlos, CA, USA.

Dziubanska, P.J., Derewenda, U., Ellena, J.F., Engel, D.A., Derewenda, Z.S., 2014. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. Acta Crystallogr. D Biol. Crystallogr. 70, 2420–2429.

Feldmann, H., Geisbert, T.W., 2011. Ebola haemorrhagic fever. Lancet 377, 849–862.

Ferreira, R.S., Simeonov, A., Jadhav, A., Eidam, O., Mott, B.T., Keiser, M.J., McKerrow, J.H., Maloney, D.J., Irwin, J.J., Shoichet, B.K., 2010. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. J. Med. Chem. 53, 4891–4905.

Geisbert, T.W., 2017. First Ebola virus vaccine to protect human beings? Lancet 389, 479–480.

Geng, C., Narasimhan, S., Rodrigues, J.P., Bonvin, A.M., 2017. Information-driven, ensemble flexible peptide docking using HADDOCK. Methods Mol. Biol. 1561, 109–138.

Group, P.I.W., Multi-National, P.I.I.S.T., Davey Jr., R.T., Dodd, L., Proschan, M.A., Neaton, J., Neuhaus Nordwall, J., Koopmeiners, J.S., Beigel, J., Tierney, J., Lane, H.C., Fauci, A.S., Massaquoi, M.B.F., Sahr, F., Malvy, D., 2016. A randomized, controlled trial of ZMapp for ebola virus infection. N. Engl. J. Med. 375, 1448–1456.

Hubbard, R.E., 2016. The role of fragment-based discovery in lead finding. In: Daniel, A., Erlanson, W.J. (Eds.), Fragment-based Drug Discovery Lessons and Outlook. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 3–36.

Hwang, T.L., Shaka, A.J., 1995. Water suppression that works - excitation sculpting using arbitrary wave-forms and pulsed-field gradients. J. Magn. Reson., Ser. A 112, 275–279.

Irwin, J.J., Duan, D., Torosyan, H., Doak, A.K., Ziebart, K.T., Sterling, T., Tumanian, G., Shoichet, B.K., 2015. An aggregation advisor for ligand discovery. J. Med. Chem. 58, 7076–7087.

Kozakov, D., Grove, L.E., Hall, D.R., Bohnuud, T., Mottarella, S.E., Luo, L., Xia, B., Beglov, D., Vajda, S., 2015. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. Nat. Protoc. 10, 733–755.

LaPlante, S.R., Aubry, N., Bolger, G., Bonneau, P., Carson, R., Coulombe, R., Sturino, C., Beaulieu, P.L., 2013. Monitoring drug self-aggregation and potential for promiscuity in off-target in vitro pharmacology screens by a practical NMR strategy. J. Med. Chem. 56, 7073–7083.

Leung, D.W., Borek, D., Luthra, P., Binning, J.M., Anantpadma, M., Liu, G., Harvey, I.B., Su, Z., Endlich-Frazier, A., Pan, J., Shabman, R.S., Chiu, W., Davey, R.A., Otwinowski, Z., Basler, C.F., Amarasinghe, G.K., 2015. An intrinsically disordered peptide from ebola virus VP35 controls viral RNA synthesis by modulating nucleoprotein-RNA interactions. Cell Rep. 11, 376–389.

Licata, J.M., Johnson, R.F., Han, Z., Harty, R.N., 2004. Contribution of ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. J. Virol. 78, 7344–7351.

Messaoudi, I., Amarasinghe, G.K., Basler, C.F., 2015. Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. Nat. Rev. Microbiol. 13, 663–676.

Muhlberger, E., 2007. Filovirus replication and transcription. Future Virol. 2, 205–215.

Muhlberger, E., Weik, M., Volchkov, V.E., Klenk, H.D., Becker, S., 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. J. Virol. 73, 2333–2342.

Noda, T., Watanabe, S., Sagara, H., Kawaoka, Y., 2007. Mapping of the VP40-binding regions of the nucleoprotein of Ebola virus. J. Virol. 81, 3554–3562.

Ruigrok, R.W., Crepin, T., Kolakofsky, D., 2011. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr. Opin. Microbiol. 14, 504–510.

Sherwood, L.J., Hayhurst, A., 2013. Ebolavirus nucleoprotein C-termini potently attract single domain antibodies enabling monoclonal affinity reagent sandwich assay (MARSA) formulation. PloS One 8, e61232.

Stark, J.L., Eghbalnia, H.R., Lee, W., Westler, W.M., Markley, J.L., 2016. NMRmix: a tool for the optimization of compound mixtures in 1D (1)H NMR ligand affinity screens. J. Proteome Res. 15, 1360–1368.

Su, Z., Wu, C., Shi, L., Luthra, P., Pintilie, G.D., Johnson, B., Porter, J.R., Ge, P., Chen, M., Liu, G., Frederick, T.E., Binning, J.M., Bowman, G.R., Zhou, Z.H., Basler, C.F., Gross, M.L., Leung, D.W., Chiu, W., Amarasinghe, G.K., 2018. Electron cryo-microscopy structure of ebola virus nucleoprotein reveals a mechanism for nucleocapsid-like assembly. Cell 172, 966–978 e912.

van Zundert, G.C.P., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P.L., Karaca, E., Melquiond, A.S.J., van Dijk, M., de Vries, S.J., Bonvin, A., 2016. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. J. Mol. Biol. 428, 720–725.

Watanabe, S., Noda, T., Kawaoka, Y., 2006. Functional mapping of the nucleoprotein of Ebola virus. J. Virol. 80, 3743–3751.

Xu, W., Luthra, P., Wu, C., Batra, J., Leung, D.W., Basler, C.F., Amarasinghe, G.K., 2017. Ebola virus VP30 and nucleoprotein interactions modulate viral RNA synthesis. Nat. Commun. 8, 15576.

Zinzalla, G., Thurston, D.E., 2009. Targeting protein-protein interactions for therapeutic intervention: a challenge for the future. Future Med. Chem. 1, 65–93.