C-Terminal Auto-Regulatory Motif of Hepatitis C Virus NS5B Interacts with Human VAPB-MSP to Form a Dynamic Replication Complex

Garvita Gupta, Jianxing Song*

Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore, Singapore

* dbssjx@nus.edu.sg

Abstract

Hepatitis C virus (HCV) is a pathogen of global importance and nearly 200 million people are chronically infected with HCV. HCV is an enveloped single-stranded RNA virus, which is characteristic of the formation of the host membrane associated replication complex. Previous functional studies have already established that the human ER-anchored VAPB protein acts as a host factor to form a complex with HCV NS5A and NS5B, which may be established as a drug target. However, there is lacking of biophysical characterization of the structures and interfaces of the complex, partly due to the dynamic nature of the complex formation and dissociation, which is extensively involved in intrinsically-disordered domains. Here by an integrated use of domain dissection and NMR spectroscopy, for the first time we have successfully deciphered that the HCV NS5B utilizes its auto-regulatory C-linker to bind the VAPB-MSP domain to form a dynamic complex. This finding implies that the NS5B C-linker is capable of playing dual roles by a switch between the folded and disordered states. Interestingly, our previous and present studies together reveal that both HCV NS5A and NS5B bind to the MSP domains of the dimeric VAP with significantly overlapped interfaces and similar affinities. The identification that EphA2 and EphA5 bind to the MSP domain with higher affinity than EphA4 provides a biophysical basis for further exploring whether other than inducing ALS-like syndrome, the HCV infection might also trigger pathogenesis associated with signalling pathways mediated by EphA2 and EphA5.

Introduction

Hepatitis C virus (HCV), first discovered in 1989 [1], is a pathogen of global importance due to significant health problems throughout the world [1–3]. Nearly 200 million people are chronically infected with HCV who are at risk of developing liver disease including liver cirrhosis and hepatocellular carcinoma. HCV is an enveloped single-stranded RNA virus, which belongs to the *Hepacivirus* genus in the *Flaviridae* family. The genome of HCV is positive sense

single ~9.6 kb long RNA and encodes ~3000 residue polyprotein, which is subsequently processed by viral and cellular proteases into 10 mature structural and nonstructural regulatory proteins on the rough endoplasmic reticulum (ER). Nonstructural proteins, which coordinate the intracellular processes of the virus life cycle, include P7 ion channel, NS2 protease, NS3 serine protease and helicase, NS4A, NS4B, NS5A proteins and NS5B RNA-dependent RNA polymerase (RdRp) [4].

HCV infection induces distinct alteration of membranes which form sponge-like inclusions, called as "membranous webs". Studies have suggested that these membranous webs constitute the sites for RNA replication, on which replication proteins and viral RNA localize [5–7]. On the membrane-associated replication machinery, RNA synthesis is catalysed by the viral RdRp activity of NS5B facilitated by both viral NS proteins and human cell host factors [8, 9]. Although viral NS proteins, RNA are main components of replication complex, the exact composition and details of replication machinery is poorly understood. Like all viruses, HCV infection acts to reprogram the cellular metabolism in such a way that infected cells devote themselves to orchestrating the production of new viruses. During this process, HCV heavily relies on human cell host factors for its replication. Thus, there is of significant interest in identifying those host factors which may be established as critical targets for drug design.

Recently, NS5B, the core component of HCV replication complex, has emerged as a key target for the development of small molecules that can selectively inhibit the replication as mammalian cells lack an RdRp equivalent. NS5B contains 591 residues, in which the catalytic domain over residues 1–531 is connected to a membrane inserting sequence 571–591 with a linker over residues 531–570 (Fig 1A). In 1999, the structures of NS5B have been determined, whose catalytic domain is composed of three subdomains (Fig 1B): palm domain which harbors the active site, finger domain and thumb domain [10–12]. Further determination of the NS5B structures together with C-linker deciphered that unlike other RdRp, the HCV NS5B has a close conformation of the active site [13–15], in which the C-linker folds back into the active center and thus regulates the enzymatic activity (Fig 1B). This unique feature appears to play a key role in RNA synthesis initiation [12–16]. Very recently, the structural changes of NS5B required for the HCV RNA replication has been successfully delineated: the thumb domain β-loop and C-linker, which is buried within the encircled active-site cavity (apo-state), vacate the active-site cavity and generate a larger cavity (open form) for RNA elongation [13–15]. As such, the C-linker acts as an auto-regulatory motif for the HCV NS5B polymerase [14–19].

Due to a high rate of mutation for viral proteins during the life cycle, targeting host-viral protein interaction interfaces for development of antiviral drugs has become increasingly attractive in recent years. Previous functional studies have identified the involvement of human vesicle-associated membrane protein-associated proteins (VAPs) in HCV replication through interaction with NS5A and NS5B [20–23]. Briefly, human VAPB protein has been biologically established by two groups to interact with both HCV NS5A and NS5B to form the HCV replication complex [20,22,23], but the high-resolution structure remains undetermined for the NS5B-VAPB interaction.

Therefore, delineation of high-resolution structures and interfaces of the VAPs-NS5A/NS5B interactions represents an essential step to facilitate further drug development by targeting their interfaces. The VAP family includes VAPA, VAPB and VAPC, out of which VAPB is composed of three domains (Fig 1C): an N-terminal nematode major sperm protein (MSP) domain, a central coiled-coil domain (CCD) and a C-terminal transmembrane domain (24–26). Previously, by high-resolution NMR characterization of structures and interactions of the differentially-dissected domains of HCV NS5A and human VAPB proteins, we have successfully deciphered that the intrinsically unstructured domain 3 of HCV NS5A forms a "fuzzy" or dynamic complex with the MSP domain of VAPB [26]. Interestingly, the mutation sites within
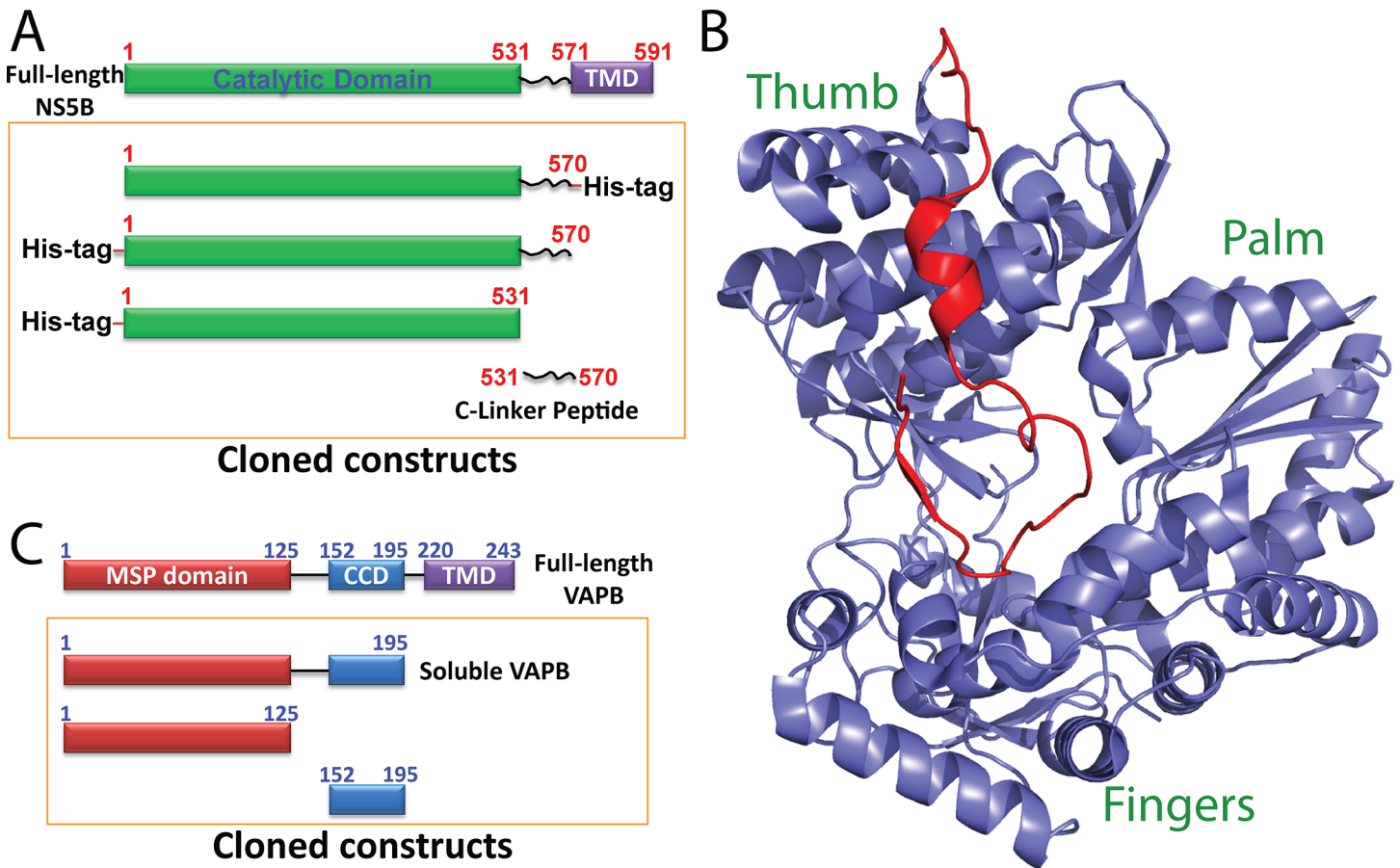
Fig 1. Domain organization and structure of HCV NS5B and human VAPB. A. Domain organization of HCV NS5B and differentially dissected proteins used in the present study. B. Crystal structure of NS5B(1–570) (PDB code of 1C2P) with the C-linker over residues 531–570 colored in red. C. Domain organization of human VAPB and differentially dissected proteins used in the present study.

doi:10.1371/journal.pone.0147278.g001

the MSP domain associated with amyotrophic lateral sclerosis (ALS) have been shown to located on the NS5A-MSP interaction interface [25–27]. On the other hand, the MSP domain has been identified to be an endogenous ligand for Eph receptors [27,28], which consists of 16 members and represents the largest transmembrane tyrosine kinase receptors [29]. Strikingly, we revealed that the interface on the MSP domain for interacting with NS5A is also overlapped with that used for binding to EphA4 [26], the only known ALS modifier [30,31], thus rationalizing a potential linkage between HCV infection and ALS pathogenesis [26,32].

In the present study, by the same approach [26,33], we aimed to NMR characterize and quantify the interactions of human VAPB with HCV NS5B as well as 6 Eph receptors. Remarkably, the results reveal that the isolated NS5B C-linker is intrinsically disordered in solution. Nevertheless, it is capable of binding the human VAPB-MSP domain over surfaces which have significant overlaps with those for binding HCV NS5A and human Eph receptors.

Results

Generation and characterization of differently dissected NS5B proteins

To study the interaction between NS5B and VAPB, we have dissected both VAPB and NS5B into different combinations of domains. For human VAPB, here we used the same dissected
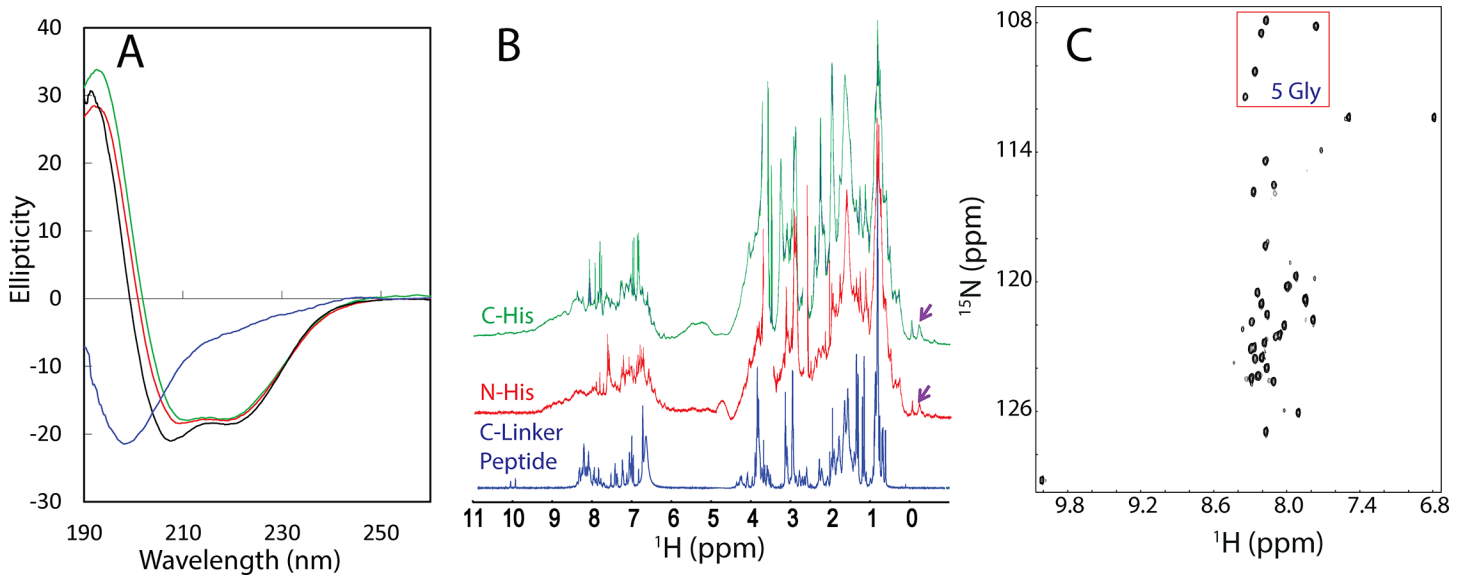
Fig 2. CD and NMR characterization of HCV NS5B proteins. A. Far UV CD spectra of C-His-tagged NS5B(1–570) (green), N-His-tagged NS5B(1–570) (red), N-His-tagged NS5B(1–531) and NS5B C-linker (531–570) (blue). B. One-dimensional ^{1}H NMR spectra of NS5B proteins. Purple arrows are used to indicate several up-field NMR peaks characteristic of a well-folded protein. C. ^{1}H-^{15}N HMR HSQC spectrum of the ^{15}N-labeled NS5B C-linker (531–570). The red box is used to indicate HSQC peaks of 5 Gly residues in the C-linker peptide.

doi:10.1371/journal.pone.0147278.g002

proteins which we previously cloned and characterized for mapping the NS5A-VAPB interactions [26], namely the soluble VAPB only with the transmembrane domain (TMD) deleted, the MSP domain and coiled-coil domain (CCD) (Fig 1C). For HCV NS5B, as it has been previously demonstrated that the removal of C-terminal membrane-interacting residues (571–591) is required to enhance the solubility but retains the similar activity of polymerase [12–14], here we cloned and expressed four dissected proteins: namely NS5B(1–570) with either N- or C-terminal His-tags, NS5B(1–531) only containing the catalytic domain [11], as well as the C-linker peptide containing residues 531–570 (Fig 1A).

All VAPB recombinant proteins were expressed and purified by the same procedures as previously described [26]. Three recombined NS5B proteins with His-tag were successfully expressed in BL21 cells and subsequently purified with Ni^{2+}-affinity chromatography followed by Heparin affinity chromatography. On the other hand, the C-linker peptide was cloned into pGEX-4T1 vector with GST-tag. The GST-fused peptide was first purified by affinity chromatography with glutathione-Sepharose 4B beads, and subsequently cleaved to release the C-linker peptide which was purified by reverse-phase (RP) C18 column using HPLC system.

As shown in Fig 2A, the NS5B(1–570) proteins with the N- and C-terminal His-tag respectively have the similar far-UV CD spectra characteristic of a well-folded helix-dominant protein, with two negative signals at ~209 and 222 nm, as well as a very large positive signal at ~193 nm, consistent with the well-defined crystal structures previously determined [10–15]. The NS5B(1–531) protein, despite having some minor difference from those of NS5B(1–570) proteins, also has a far-UV CD spectra characteristic of a well-folded helix-dominant protein. In fact, comparison of the crystal structures of NS5B(1–570) [15] and NS5B(1–531) [10,11] reveals that the catalytic domain have almost the same structures in two constructs. Consistent with CD results, their ^{1}H one-dimensional NMR spectra of NS5B(1–570) proteins are also similar, both of which have several very up-field peaks (Fig 2B), clearly suggesting that they are well-folded with the tight tertiary packing. Unfortunately, due to the short transverse relaxation time (T2) resulting from the very large molecular size, the NMR resonance peaks are very

broad particularly over the amide proton region (Fig 2B). Consequently, no high-quality ^{1}H-^{15}N NMR HSQC spectrum could be acquired.

By contrast, as shown in Fig 2A, the C-linker peptide has a far-UV CD spectrum typical of a highly-disordered protein with the maximal negative signal at 198 nm, and lacking any positive signal at 190 nm. Furthermore, very up-field peaks are totally absent in the ^{1}H one-dimensional NMR spectrum (Fig 2B), clearly indicative of the lack of any tight tertiary packing. Its ^{1}H-^{15}N NMR HSQC spectrum (Fig 2C) also indicated that it lacks any tight tertiary packing as judged from very narrow ^{1}H (~0.86 ppm) and ^{15}N (~19 ppm) spectral dispersions. More specifically, the peptide contains 40 residues plus two extra N-terminal Gly-Ser residues from the cleavage of the GST-fusion protein, in which there are 6 Gly and 3 Pro residues. As the HSQC peak of the N-terminal Gly is usually undetectable due to the fast exchange of its amide protons with water, and 3 Pro residues have no amide proteins, there should be 38 backbone HSQC peaks in the HSQC spectrum. However, although HSQC peaks of 5 Gly residues within the NS5B C-linker peptide are all detectable and well-separated (Fig 2C), only ~32 strong HSQC peaks could be detected for backbone amides (Fig 2C). This observation implies that there are 6 residues which might have their HSQC peaks overlapped with other; or/and become very weak or even undetectable due to their fast exchange with water, or/and due to conformational exchanges on μs-ms time scale. Therefore, as judged by its CD, 1D and HSQC spectra, the isolated C-linker appears to be predominantly disordered in solution, lacking of any well-formed secondary structures and tight tertiary packing. Unfortunately, due to its aggregation at high concentration, we attempted but failed to collect high-quality triple-resonance NMR experiments to achieve its assignments.

Identification of the binding interfaces on VAPB and NS5B

Previously, we have cloned and shown that the NS5B(1–570) with a C-terminal His-tag was activity in binding the human VAPC protein [33] and DDX5 [34]. So here we first tested its binding to the soluble full-length VAPB. However, the NMR titrations indicated a lack of significant interaction between them, disagreeing with the previous functional results that HCV NS5B interacts with human VAPB protein to form the replication complex [20,22,23]. This led us to suspecting that the binding region of NS5B to VAPB is mostly located on the C-terminus and therefore the presence of the C-terminal His-tag would block the binding, possibly because the His-tag at the C-terminus may be incompatible with interacting with the MSP domain, or/and alter the dynamics of the C-linker because a comparison of the NS5B (1–570) structures with the His-tag at the N-terminus (PDB code of 1C2P) [11] or at the C-terminus (PDB code of 1QUV) [12] indicates that the C-terminal linker adopts a similar structure in both crystal structures. Therefore, we then cloned the NS5B(1–570) with a N-terminal His-tag and characterized its binding to the soluble VAPB protein. Indeed, as shown in Fig 3A, the addition of the unlabelled N-His-tagged NS5B(1–570) induced a disappearance of many HSQC peaks of the ^{15}N-labeled VAPB, clearly indicating the significant binding between two proteins. With NMR assignments of the VAPB domains we previously published [25–27], we thus identified that the disappeared peaks are almost all from the MSP domain of VAPB, while the remaining peaks (red) are from the coiled-coil domain (CCD) of VAPB. We also titrated the ^{15}N-labeled CCD of VAPB by both NS5B(1–570) proteins but found no significant perturbation. These results clearly reveal that HCV NS5B binds to the MSP domain of human VAPB protein, completely consistent with previous biological studies [20,23,24].

To identify the region of HCV NS5B which binds the MSP domain, we conducted detailed HSQC titrations of the ^{15}N-labeled MSP domain by different NS5B proteins. The C-His-tagged NS5B(1–570) only induced slight shifts of several HSQC peaks of the MSP domain (Fig 3B),
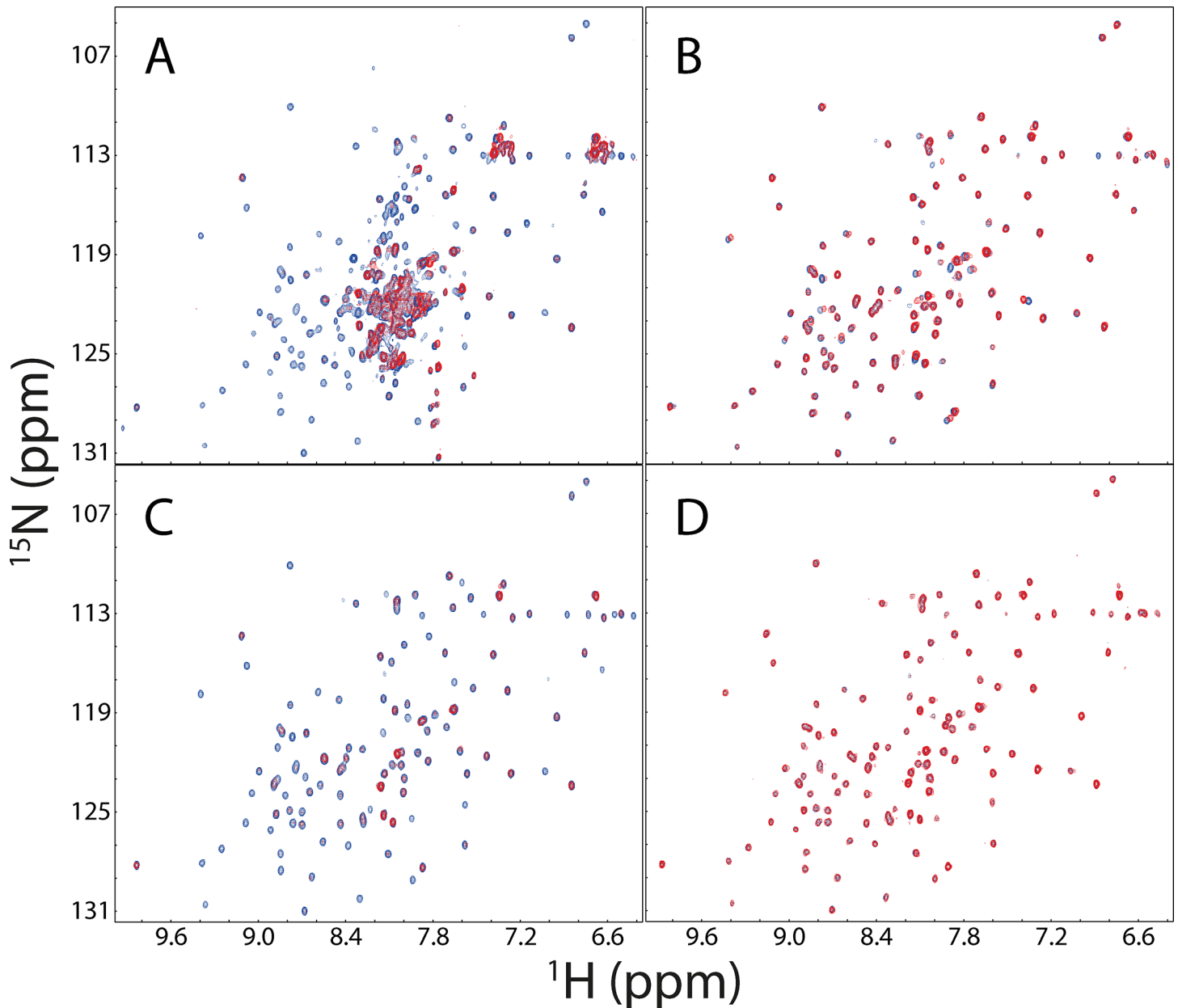
Fig 3. HSQC characterization of the binding between VAPB and NS5B. A. ^{1}H-^{15}N HSQC spectra of the ^{15}N-labeled soluble VAPB protein (1–195) in the absence (blue) and in the presence of the unlabelled N-His-tagged NS5B(1–570) at a molar ratio of 1:2 (red). In the presence of the unlabelled N-His-tagged NS5B(1–570), most well-dispersed peaks disappeared which were identified to be from the MSP domain while the remaining peaks (red) are from the coiled coil domain, based on our NMR assignments of the human VAPB domains previously published (25–27). B. ^{1}H-^{15}N HSQC spectra of the ^{15}N-labeled VAPB-MSP domain in the absence (blue) and in the presence of the unlabelled C-His-tagged NS5B(1–570) at a molar ratio of 1:4 (red). C. ^{1}H-^{15}N HSQC spectra of the ^{15}N-labeled MSP domain in the absence (blue) and in the presence of the unlabelled N-His-tagged NS5B(1–570) at a molar ratio of 1:2 (red). D. ^{1}H-^{15}N HSQC spectra of the ^{15}N-labeled MSP domain in the absence (blue) and in the presence of the unlabelled N-His-tagged NS5B(1–531) at a molar ratio of 1:4 (red).

doi:10.1371/journal.pone.0147278.g003

while the N-His-tagged NS5B(1–570) triggered a dramatic disappearance of the majority of the peaks (Fig 3C). Most strikingly, the NS5B(1–531) containing only the catalytic domain shows almost no perturbation on the HSQC peaks of the MSP domain. These results together
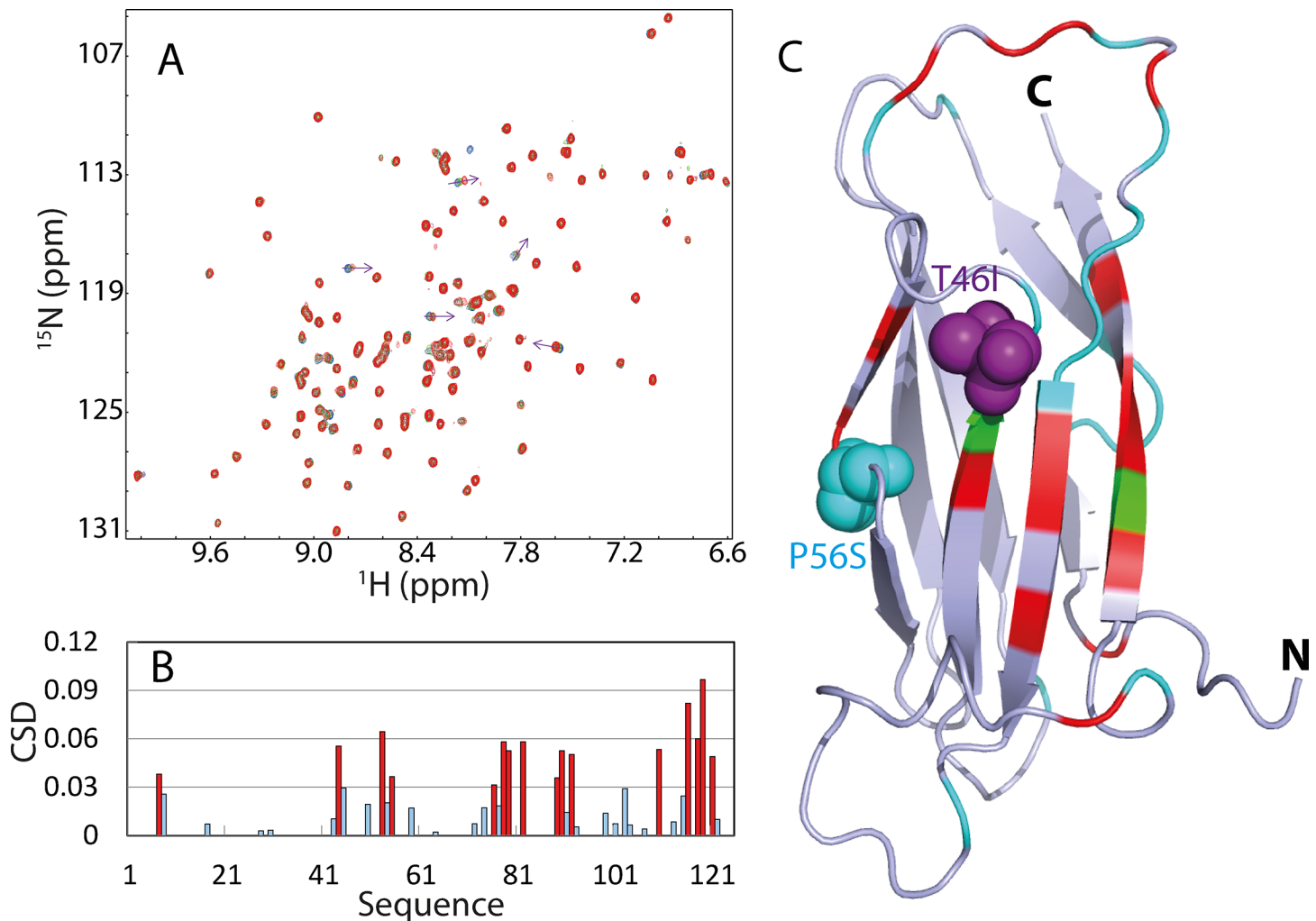
Fig 4. Quantification of the binding between MSP domain and NS5B C-linker. A. ^1H-^{15}N HSQC spectra of the ^{15}N-labeled MSP domain in the absence (blue) and in the presence of the unlabelled NS5B C-linker (531–570) at a molar ratio of 1:1 (green) and 1:3 (red). B. Residue-specific chemical shift difference (CSD) of the MSP domain in the presence of the NS5B C-linker (531–570) at a molar ratio of 1:3. The bars for the residues with the CSD value > 0.03 (average + STD) are colored in red. C. Crystal structure of the MSP domain we previously determined in which cyan is used for indicating Pro and residues with missing HSQC peaks in the free state; green for residues disappeared in the presence of the NS5B C-linker (531–570), and red for the residues with the CSD value > 0.03 at a molar ratio of 1:3. The pink and cyan spheres are used to indicate the sites of the mutations Thr46 (to Ile) and Pro56 (to Ser) respectively, which are associated with familial ALS.

doi:10.1371/journal.pone.0147278.g004

decipher that the C-linker of the HCV NS5B plays a central role in binding the MSP domain of the human VAPB.

Quantification of the binding between NS5B C-linker and MSP domain

We thus generated the recombinant NS5B C-linker peptide which is highly disordered in solution (Fig 2). However, it is active in binding the MSP domain. As shown in Fig 4A, a gradual addition of the peptide induces both shifts and intensity reductions of a set of HSQC peaks of the ^{15}N-labeled MSP domain. As the peptide started to show visible aggregation at high concentration, we were only able to titrate the MSP domain with the peptide at the molar ratio of 1:3 (MSP:peptide) at which the shifts of the peaks are almost saturated. However, the C-linker peptide alone appears to have less ability to provoke the disappearance of the MSP peaks than

Table 1. Dissociation constants for the binding of the MSP domain with NS5B C-linker, EphA2 and EphA5 obtained by fitting NMR titration data.

Residue	NS5B C-Tail Kd (µM) ± SE	EphA2 Kd (µM) ± SE	EphA5 Kd (µM) ± SE
6Q		7.9 ± 0.6	
44V		4.1 ± 1.2	36.2 ± 2.8
53C	1.4 ± 0.3	6.3 ± 0.3	57.1 ± 0.2
76F			59.2 ± 3.4
78Y		10.1 ± 1.1	89.2 ± 8.9
79D	13.6 ± 1.1	18.9 ±1.6	77.9 ± 4.2
82E	12.8 ± 0.9		
90V	9.0 ± 1.2		53.8 ± 1.1
91Q			74.6 ± 5.0
92S			81.5 ± 4.9
116D	5.7 ± 0.7		
118K	2.9 ± 0.6		
119L	16.0 ± 2.8		
122V		14.2 ± 0.7	
Average	**8.8 ± 1.1**	**10.3 ± 0.9**	**66.2 ± 3.8**

doi:10.1371/journal.pone.0147278.t001

the N-His-tagged NS5B(1–570), probably because the presence of the very large catalytic domain might provide additional binding residues, or/and trigger radical dynamics of the MSP domain on µs-ms time scale.

Fig 4B presents the chemical shift difference (CSD) of the HSQC peaks for the MSP domain at 1:3 (MSP:peptide), while Fig 4C shows the MSP structure we previously determined [25], whose residues are colored if their HSQC peaks disappeared (Thr46, Ser117), or significantly shifted with the CSD value > 0.03 (average + STD). Interestingly, these residues are all located on one side of the MSP structure (Fig 4C). Most strikingly, two ALS-causing mutations Pro56Ser and Thr46Ile are also located on this side and in particular the HSQC peaks of Thr46 disappeared upon binding to the NS5B C-linker. Remarkably, the binding interface of the MSP domain to the NS5B C-linker is highly overlapped with that for binding the NS5A Domain 3 as we previously identified by NMR [26].

With the well-established procedure to obtain the dissociation constants (Kd) by fitting the data of HSQC titrations, which we have previously utilized to quantify the NS5A-MSP [26] and NS5B-VAPC [33] interactions, here we gain the residue-specific Kd values for significantly-shifted residues, with an average of 8.8 µM (Table 1). It is worth to note that this value is highly similar to that (8.0 µM) for the binding of the MSP domain to the HCV NS5A D3 domain [26]. The results here hence reveal that the HCV NS5A and NS5B bind to the highly overlapped interfaces of the human MSP domain with very similar affinities.

Characterization of the binding of the MSP domain with Eph receptors

Previously, the human MSP domain has been biologically identified to be cleaved and subsequently secreted to serve as a ligand of the ligand binding domain (LBD) of Eph receptors, and furthermore it was shown that the interaction between EphA4 and VAPB-MSP domain is involved in ALS [28,30,31]. However, so far the high-resolution biophysical characterization has been conducted only on the MSP-EphA4 interaction previously by us [26,27]. Here we further extended to characterizing the interactions of the MSP domain to EphA2, EphA5, EphA6,
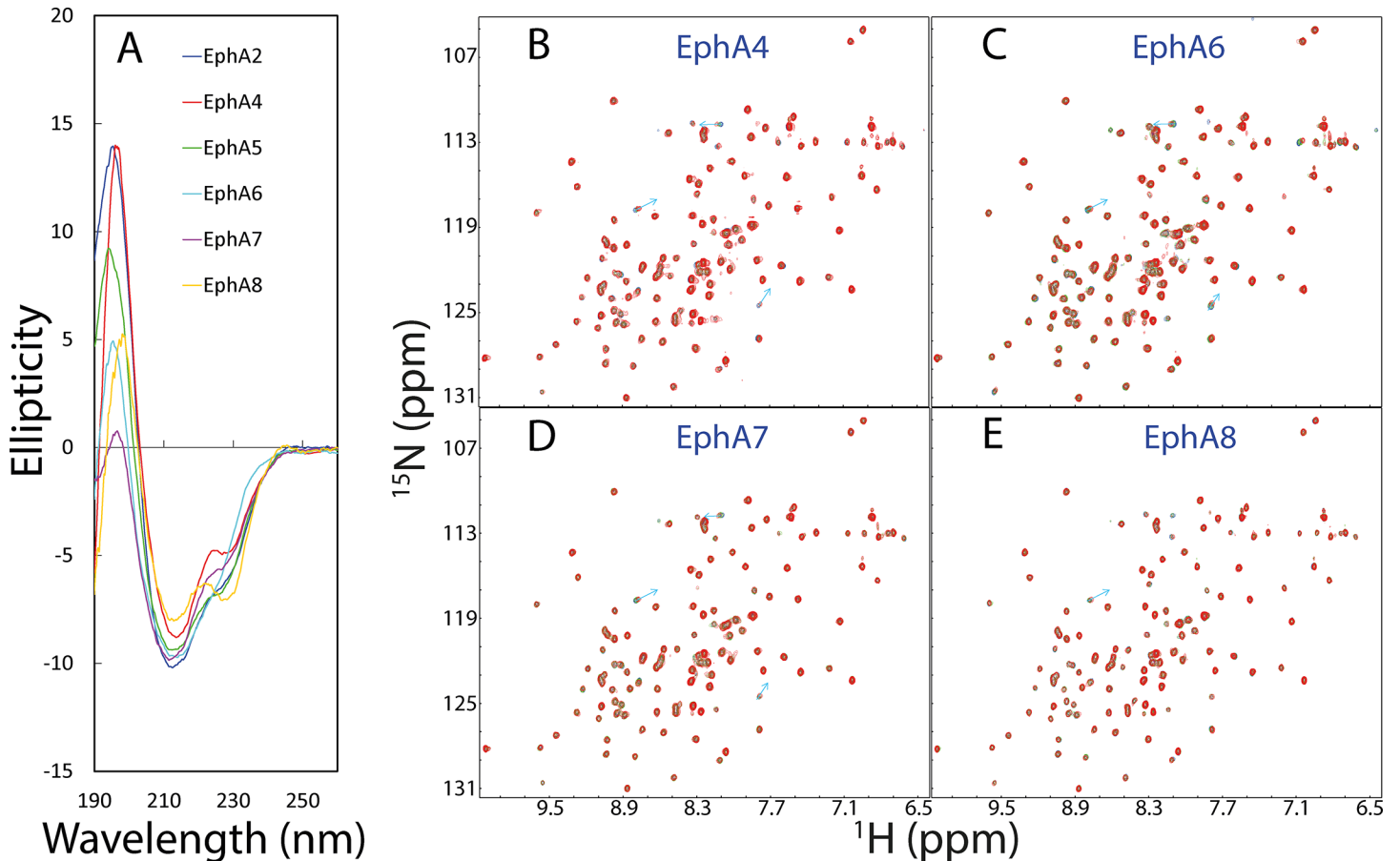
Fig 5. Characterization of the binding between MSP and Eph LBDs. A. Far-UV CD spectra of 6 Eph LBDs. ^{1}H-^{15}N HSQC spectra of the ^{15}N-labeled MSP domain in the absence (blue) and in the presence of the unlabelled LBDs of EphA4 (B), EphA6 (C), EphA7 (D) and EphA8 (E) at a molar ratio of 1:1 (green) and 1:4 (red).

doi:10.1371/journal.pone.0147278.g005

EphA7 and EphA8. These Eph LBDs were expressed and purified by Ni^{2+}-NTA affinity chromatography followed by FPLC size exclusion chromatography as we previously described [26,27,31,35–38].

As shown in Fig 5A, all 6 Eph LBDs are well-folded as judged from their far-UV CD spectra. Subsequently, we titrated the ^{15}N-labeled MSP domain with the unlabelled Eph LBDs. As previously we have extensively characterized the interaction of the MSP domain with EphA4 LBD by ITC, NMR and modelling [27], here we used it as a reference. Exactly as we have previously shown [27], only very minor shifts of several HSQC peaks could be observed for the ^{15}N-labeled MSP domain (Fig 5B), upon adding the unlabelled EphA4 LBD at a molar ratio of 1:4 (MSP:Eph4). Similar scenarios are also observed for EphA6 (Fig 5C), EphA7 (Fig 5D) and EphA8 (Fig 5E). This implies that the binding affinities for the MSP domain to EphA6, EphA7 and EphA8 are similar to or even weaker than that to EphA4.

On the other hand, the HSQC titrations reveal that EphA2 (Fig 6A) and EphA5 (Fig 6B) LBDs are able to trigger both significant shift and intensity reduction for many HSQC peaks of the ^{15}N-labeled MSP domain at low molar ratios, and particularly the peak shifts have become mostly saturated for EphA2 LBD at 1:4 (MSP:EphA2) and for EphA5 at 1:6 (MSP:EphA2). Fig 6C presents the chemical shift differences (CSD) of the MSP HSQC peaks in the presence of EphA2 and EphA5 at 1:4 (MSP:Eph), while Fig 6D and 6E show the MSP structures, whose
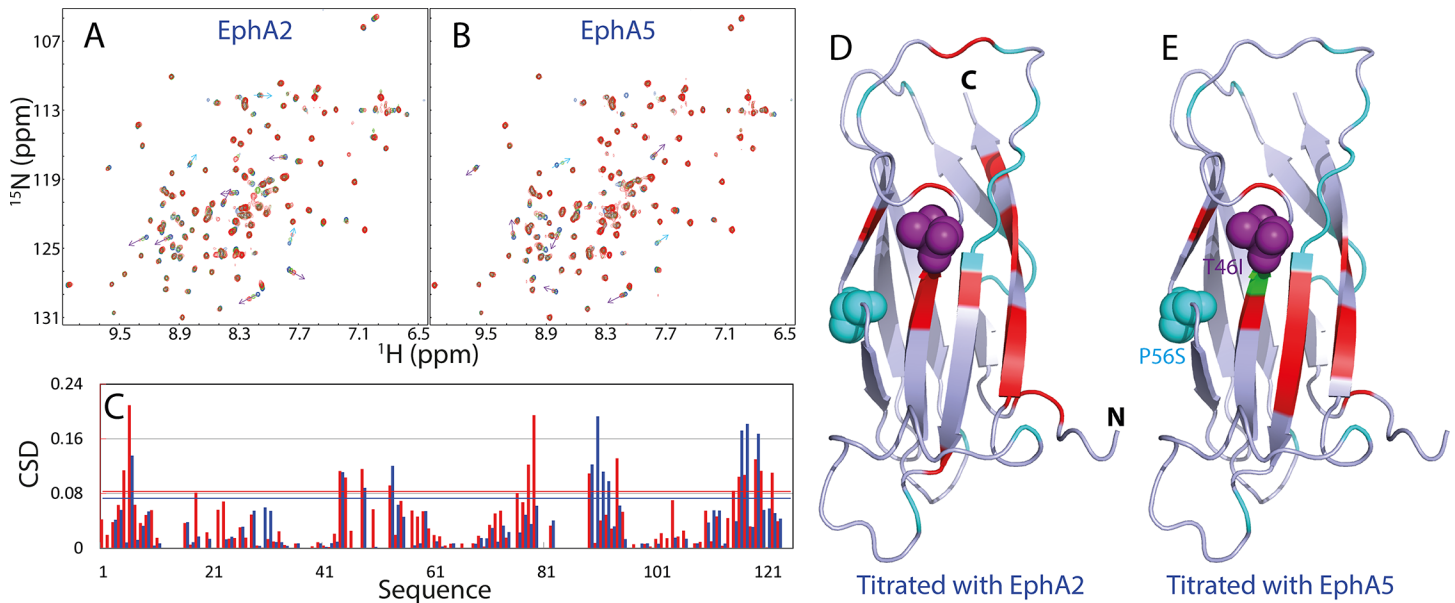
Fig 6. Quantification of the binding between MSP domain and EphA2/EphA5. ^1H-^{15}N HSQC spectra of the ^{15}N-labeled MSP domain in the absence (blue) and in the presence of the unlabelled LBDs of EphA2 (A) and EphA5 (B) at a molar ratio of 1:1 (green) and 1:3 (red). C. Residue-specific chemical shift difference (CSD) of the MSP domain in the presence of EphA2 (red) and EphA5 (blue) at a molar ratio of 1:4. Crystal structure of the MSP domain in which cyan is used for indicating Pro and residues with missing HSQC peaks in the free state; green for residues disappeared in the presence of EphA2 (D) or EphA5 (E), and red for the residues with the CSD value > 0.083 for EphA2 and >0.073 for EphA5 at a molar ratio of 1:4. The pink and cyan spheres are used to indicate the sites of the mutations Thr46 (to Ile) and Pro56 (to Ser) respectively, which are associated with familial ALS.

doi:10.1371/journal.pone.0147278.g006

residues are colored if their HSQC peaks disappeared, or significantly shifted with the CSD value > 0.083 for EphA2 and 0.073 for EphA5 (average + STD). Interestingly, the significantly perturbed residues triggered by binding to EphA2 and EphA5 are located on the same side of the MSP structure (Fig 4C), which is also highly similar to those by the NS5B C-linker (Fig 4C), and by EphA4 [27].

We also obtained the Kd values for significantly shifted residues induced by binding with EphA2 and EphA5, which are presented in Table 1. The average Kd value is 10.3 µM for the MSP-EphA2 interaction while is 66.2 µM for the MSP-EphA5. This suggests that EphA2 LBD binds to the MSP with the highest affinity among 6 Eph LBDs investigated here. In the future, we are also interested in characterizing the binding affinity of the MSP domain with the rest of Eph receptors.

Discussion

RNA viruses such as HCV are characteristic of the formation of the host membrane associated replication complexes [20–23,39,40]. Although it is widely anticipated that studies of RNA virus replication machineries will have a critical impact on antiviral drug development because of their essential and specific roles in virus replication, currently high-resolution structures and interaction interfaces remain relatively unknown for the HCV replication complexes [41]. Previously, it has been biologically established that both HCV NS5A and NS5B interact with the human VAPB proteins to constitute the core components of the HCV replication complex [20–23] but so far there had been lacking of extensive atomic-resolution structural and binding characterization. One challenge for the biophysical studies is associated with the dynamic nature of the complex formation and dissociation, with extensive involvement of intrinsically disordered domains/regions. Although the dynamics of the replication complex play a central

role in implementing its functions, this makes it very challenging for crystallographic investigations. In this regard, NMR spectroscopy represents the only powerful tool to delineate the weak binding involved in intrinsically disordered domains at atomic resolution [26, 42–48]. Indeed, previously we have conducted a systematic investigation on the NS5A-VAPB interactions by dissecting both VAPB and NS5A proteins into 14 domains/fragments, followed by extensive structural and binding characterizations with heteronuclear NMR spectroscopy. Our study successfully reveals that the intrinsically disordered Domain 3 of NS5A is responsible for dynamically binding to the MSP domain of VAPB [26].

In the present study, by use of the domain dissection followed by NMR characterization, for the first time we have successfully deciphered that the HCV NS5B binds to the MSP domain of the human VAPB mostly by its C-linker. This finding is interesting as in the NS5B(1–570) structure with a close conformation of the active site [13–15], the C-linker folds back into the active center of the catalytic domain (Fig 1B), which thus acts to auto-regulate the RdRp enzymatic activity [13–15,49]. The regulatory role of the C-linker in mediating catalysis is also supported by previous functional studies, which showed that the enzymatic activity of the C-terminally truncated NS5B(1–531) could be up to ~150-fold higher than that of NS5B(1–570) [13,14,49–51]. Interestingly, the C-linker could be clearly visualized in crystal structures [13–15], which is composed of a short helix and β-strand despite also having regions without regular secondary structures. As such, it is expected that in the context of NS5B(1–570) having the catalytic domain, the C-linker would have less dynamic secondary structures and tighter packing than what observed on the isolated C-linker in solution, which has no stable secondary structure and tight packing as clearly indicated by CD and NMR. As such, our study suggests that the C-linker is able to play dual roles by a conformational and dynamic transformation between the folded and disordered states, very similar to what we previously observed on the auto-inhibitory domain of PAK4 kinase [48]. Remarkably, our study also reveals that the HCV NS5B C-linker and NS5A domain 3 bind the human VAPB MSP domain, with the similar affinities and interfaces on the MSP domain [26]. The high dynamics required for formation and dissociation of this replication complex appears to be nicely achieved by the relative-weak affinity of the interactions of HCV NS5A/NS5B to the human MSP domain, which is made possible by extensive involvement of intrinsically disordered HCV domains including NS5A Domain 3 and NS5B C-linker.

Previously the MSP domain has been functionally identified to act as an endogenous inhibitor for the ephrin-binding domain of Eph receptor family [28], which is composed of 10 EphA and 6 EphB members with a large binding promiscuity [29]. On the other hand, P56S and T46I mutations disrupting the MSP fold have been shown to cause ALS [25–27,52,53], while EphA4 was identified to be the only known ALS modifier [30,31]. This implies that the binding of the MSP domain by HCV proteins will reduce the secretion of the MSP domain to function as an inhibitor of Eph receptors. This may lead to the similar syndromes triggered by other ALS pathogenesis, which include MSP mutations or/and reduced inhibition of Eph receptors. Indeed, ALS-like syndrome has been observed on the patient with chronic hepatitis C [32,54]. However, so far, only the interaction of the MSP domain with EphA4 has been biophysically characterized [27]. Here, we further characterized the interaction of the MSP domain with five more Eph receptors, namely EphA2, EphA5, EphA6, EphA7 and EphA8. The results reveal that EphA2 and EphA5 have higher affinity than EphA4, despite having the similar binding interfaces on the MSP domain. Our present results provide biophysical clues for further functional assessment about whether the HCV infection may additionally affect the signal pathways associated with EphA2 and EphA5.

In summary, for the first time we have successfully deciphered that the HCV NS5B utilizes its C-linker to bind the MSP domain of the human VAPB to form a dynamic complex. This

finding implies that NS5B C-linker achieves dual roles by a conformational/dynamic switch between the folded state acting as an enzymatic auto-inhibitor, and disordered state ready for dynamically binding the host factor VAPB to initiate RNA synthesis. The identification that EphA2 and EphA5 bind the MSP domain with higher affinity than EphA4 offers clues for further addressing whether other than inducing ALS-like syndrome, the HCV infection might also trigger pathological consequences associated with signalling pathways mediated by EphA2 and EphA5.

Materials and Methods

DNA cloning

The DNA fragments encoding the human VAPB(1–195), MSP(1–125) and CCD(152–195) domains were amplified from HeLa cell cDNA library with designed primers as previously described [22]. All the VAPB DNAs were cloned into a modified pET32a vector (Novagen). DNA fragments encoding NS5B proteins were generated from HCV genotype 1b strain S1 (GenBank-AF356827.1). DNA encoding NS5B(1–570) was cloned into pET32a and pET22b expression vectors with N- and C-terminal His-tag respectively, while NS5B(1–531) was cloned into pET32a. The DNA fragment for the C-linker over residues 531–570 was amplified and cloned into pGEX-4T1 vector.

DNA fragments encoding the ephrin binding domain of Eph receptors (EphA2, EphA3, EphA4, EphA5, EphA6, EphA7, EphA8 receptor) were cloned into pET 32a (Novagen) as we previously described [26,27,31,33–38].

Expression and purification of VAPB and NS5B proteins

After the sequences were confirmed by DNA sequencing analysis, plasmids were transformed into *Escherichia coli* BL21 (DE3) star (Invitrogen) competent cells. For the expression of recombinant protein, transformed cells with respective plasmids were grown in Luria-Bertani (LB) medium with 100 μg/ml ampicillin at 37°C to reach the absorbance of 0.6 at 600 nm and subsequently induced with respective optimized IPTG concentrations.

All the VAPB proteins were expressed and purified as we previously described [26]. For purifying NS5B proteins, harvested cells were resuspended and lysed by sonication in lysis buffer (50 mM Tris, 500 mM NaCl, 10% glycerol, 20 mM imidazole, 10 mM 2-mercaptoethanol, pH 7.5) containing protease inhibitor cocktail (Roche). His-tagged NS5B proteins were purified by Ni^{2+}-affinity chromatography (Qiagen) while GST-fused NS5B(531–570) was purified by affinity chromatography with glutathione-Sepharose 4B beads (Pharmacia Biotech) under native conditions. Further purification of His-tagged proteins was achieved with Heparin Sepharose HiTrap HP column (GE healthcare). GST-tagged NS5B(531–570) was cleaved with thrombin and then the released C-linker peptide was purified by HPLC on a reverse-phase C18 column (Vydac).

Plasmids containing the ligand binding domain of 6 Eph receptors were transformed into *Escherichia coli* Rosetta-gami (DE3) cells (Novagen). Transformed cells were grown in LB media and protein was purified under native conditions using affinity chromatography followed by thrombin cleavage to remove the His-tag. Cleaved Eph proteins were further purified with HPLC size exclusion chromatography using HiLoad 16/60 Superdex 200 column (GE healthcare).

The production of the isotope-labeled VAPB and NS5B domain/fragment proteins for NMR studies followed a similar procedure except that the bacteria were grown in M9 medium with the addition of (^{15}NH$_4$)$_2$SO$_4$ for ^{15}N-labeling, or (^{15}NH$_4$)$_2$SO$_4$/^{13}C-glucose for ^{15}N/^{13}C-labeling. The purity of all protein samples was checked by the SDS-PAGE gel and their

molecular weights were verified by a Voyager STR matrix-assisted laser desorption ionization time-of-flight-mass spectrometer (Applied Biosystems). The concentration of protein samples was determined by the spectroscopic method in the presence of denaturant [26].

CD and NMR experiments

All far-CD experiments were carried out in a Jasco J-810 spectropolarimeter (Jasco Corporation, Tokyo, Japan) at 25°C as we described previously [26] at a protein concentration of 20 μM in 2 mM phosphate buffer (pH 6.8). Data from five independent scans were added and averaged.

All NMR samples were prepared in 10 mM phosphate buffer at pH 6.8. NMR data were collected at 25°C on an 800-MHz Bruker Avance spectrometer equipped with a shielded cryoprobe. Protein samples at a concentration of 100 μM were used for ^{1}H one-dimensional or ^{1}H-^{15}N NMR two-dimensional HSQC characterization.

NMR HSQC titrations

For NMR titration experiments, two-dimensional ^{1}H-^{15}N HSQC spectra of ^{15}N- labeled VAPB proteins were acquired at a protein concentration of 50 μM in the absence or presence of the unlabeled NS5B/Eph proteins at different molar ratios in 10 mM phosphate buffer at pH 6.8. More specifically, for characterizing the interactions of the MSP domain with the NS5B C-linker peptide or human Eph receptors, a series of HSQC spectra of the ^{15}N-labeled MSP domain at a protein concentration of 50 μM were collected in 10 mM phosphate buffer at pH 6.8, with gradual addition of the C-linker peptide or Eph receptors with a ratio increment of 0.5 until the shift of HSQC peaks were mostly saturated [26]. By superimposing the HSQC spectra at different molar ratios, the shifted or disappeared HSQC peaks could be identified and further assigned to the corresponding residues based on NMR assignments of VAPB domains we previously published [25–27]. The degree of perturbation was reflected by an integrated chemical shift difference (CSD) calculated by the formula $((\Delta^1H)^2+(\Delta^{15}N)^2/5)^{1/2}$ [26,33]. The calculated CSD tracings at various different molar ratios were fitted by using one binding site model [26,33] to obtain residue-specific dissociation constants (Kd), which are summarized in Table 1.

Author Contributions

Conceived and designed the experiments: JXS GG. Performed the experiments: GG. Analyzed the data: GG JXS. Wrote the paper: GG JXS.

References

1. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 1989; 244: 359–562 PMID: 2523562

2. Thiel HJ, Collett MS, Gould EA, Heinz FX, Houghton M, Meyers G, et al. Flaviviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA, editors. Virus Taxonomy, VIIIth Report of the ICTV. 2005. pp. 979–996.

3. Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. Lancet Infect Dis. 2005; 5: 558–567 PMID: 16122679

4. Reed KE, Rice CM. Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. Curr Top Microbiol Immunol. 2000; 242: 55–84. PMID: 10592656

5. Egger D, Wolk B, Gosert R, Bianchi L, Blum HE, Moradpour D, et al. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. J Virol. 2002; 76: 5974–5984 PMID: 12021330

6. Gosert R, Egger D, Lohmann V, Bartenschlager R, Blum HE, Bienz K, et al. Identification of the hepatitis C virus RNA replication complex in huh-7 cells harboring subgenomic replicons. J Virol. 2003; 77: 5487–5492 PMID: 12692249

7. Alvisi G, Madan V, Bartenschlager R. Hepatitis C virus and host cell lipids: an intimate connection. RNA Biol. 2011; 8: 258–269. PMID: 21593584

8. Lohmann V. Hepatitis C virus RNA replication. Curr Top Microbiol Immunol. 2013; 369: 167–198. doi: 10.1007/978-3-642-27340-7_7 PMID: 23463201

9. Moradpour D, Penin F, Rice CM. Replication of hepatitis C virus. Nature Rev. Microbiol. 2007; 5: 453–463.

10. Bressanelli S, Tomei L, Roussel A, Incitti I, Vitale RL, Mathieu M, et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci U S A. 1999; 96: 13034–13039. PMID: 10557268

11. Lesburg CA, Cable MB, Ferrari E, Hong Z, Mannarino AF, Weber PC. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol. 1999; 6: 937–943. PMID: 10504728

12. Ago H, Adachi T, Yoshida A, Yamamoto M, Habuka N, Yatsunami K, et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Structure. 1999; 7: 1417–1426. PMID: 10574802

13. Simister P, Schmitt M, Geitmann M, Wicht O, Danielson UH, Klein R, et al. Structural and functional analysis of hepatitis C virus strain JFH1 polymerase J.Virol. 2009; 83: 11926–11939 doi: 10.1128/JVI. 01008-09 PMID: 19740982

14. Schmitt M, Scrima N, Radujkovic D, Caillet-Saguy C, Simister PC, Friebe P, et al. Comprehensive Structure-Function Comparison of Hepatitis C Virus Strains Jfh1 and J6 Polymerases Reveals a Key Residue Stimulating Replication in Cell Culture Across Genotypes. J.Virol. 2011; 85: 2565–2581 doi: 10.1128/JVI.02177-10 PMID: 21209117

15. Appleby TC, Perry JK, Murakami E, Barauskas O, Feng J, Cho A, et al. Viral replication. Structural basis for RNA replication by the hepatitis C virus polymerase. Science. 2015; 347: 771–775. doi: 10. 1126/science.1259210 PMID: 25678663

16. Harrus D, Ahmed-El-Sayed N, Simister PC, Miller S, Triconnet M, Hagedorn CH, et al. Further insights into the roles of GTP and the C terminus of the hepatitis C virus polymerase in the initiation of RNA synthesis. J. Biol. Chem. 2010; 285: 32906–32918. doi: 10.1074/jbc.M110.151316 PMID: 20729191

17. Lévêque VJ, Johnson RB, Parsons S, Ren J, Xie C, Zhang F, et al. Identification of a C-terminal regulatory motif in hepatitis C virus RNA-dependent RNA polymerase: structural and biochemical analysis. J Virol. 2003; 77: 9020–9028. PMID: 12885918

18. Reich S, Golbik RP, Geissler R, Lilie H, Behrens SE. Mechanisms of activity and inhibition of the hepatitis C virus RNA-dependent RNA polymerase. J Biol Chem. 2010; 285: 13685–13693. doi: 10.1074/jbc. M109.082206 PMID: 20194503

19. Wang M, Ng KK, Cherney MM, Chan L, Yannopoulos CG, Bedard J, et al. Nonnucleoside analogue inhibitors bind to an allosteric site on HCV NS5B polymerase. Crystal structures and mechanism of inhibition. J Biol Chem. 2003; 278: 9489–9495. PMID: 12509436

20. Gao L, Aizaki H, He JW, Lai MM. Interactions between viral nonstructural proteins and host protein hVAP-33 mediate the formation of hepatitis C virus RNA replication complex on lipid raft. J Virol. 2004; 78: 3480–3488. PMID: 15016871

21. Evans MJ, Rice CM, Goff SP. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and viral RNA replication. Proc Natl Acad Sci U S A. 2004; 101: 13038–12043. PMID: 15326295

22. Hamamoto I, Nishimura Y, Okamoto T, Aizaki H, Liu M, Mori Y, et al. Human VAP-B is involved in hepatitis C virus replication through interaction with NS5A and NS5B. J Virol. 2005; 79: 13473–13482. PMID: 16227268

23. Aizaki H, Lee KJ, Sung VM, Ishiko H, Lai MM. Characterization of the hepatitis C virus RNA replication complex associated with lipid rafts. Virology. 2004; 324: 450–461. PMID: 15207630

24. Skehel PA, Martin KC, Kandel ER, Bartsch D. A VAMP-binding protein from Aplysia required for neurotransmitter release. Science. 1995; 269: 1580–1583. PMID: 7667638

25. Shi J, Lua S, Tong JS, Song J. Elimination of the native structure and solubility of the hVAPB MSP domain by the Pro56Ser mutation that causes amyotrophic lateral sclerosis. Biochemistry. 2010; 49: 3887–3897. doi: 10.1021/bi902057a PMID: 20377183

26. Gupta G, Qin H, Song J. Intrinsically unstructured domain 3 of hepatitis C Virus NS5A forms a "fuzzy complex" with VAPB-MSP domain which carries ALS-causing mutations. PLoS One. 2012; 7: e39261. doi: 10.1371/journal.pone.0039261 PMID: 22720086

27. Lua S, Qin H, Lim L, Shi J, Gupta G, Song J. Structural, stability, dynamic and binding properties of the ALS-causing T46I mutant of the hVAPB MSP domain as revealed by NMR and MD simulations. PLoS One 2011; 6: e27072. doi: 10.1371/journal.pone.0027072 PMID: 22069488

28. Tsuda H, Han SM, Yang Y, Tong C, Lin YQ, Mohan K, et al. The amyotrophic lateral sclerosis 8 protein VAPB is cleaved, secreted, and acts as a ligand for Eph receptors. Cell. 2008; 133: 963–77. doi: 10.1016/j.cell.2008.04.039 PMID: 18555774

29. Pasquale BE. Eph-ephrin bidirectional signaling in physiology and disease. Cell. 2008; 133: 38–52. doi: 10.1016/j.cell.2008.03.011 PMID: 18394988

30. Van Hoecke A, Schoonaet L, Lemmens R, Timmers M, Staats KA, Laird AS, et al. EPHA4 is a disease modifier of amyotrophic lateral sclerosis in animal models and in humans. Nat. Med. 2012; 18: 1418–1422. PMID: 22922411

31. Qin H, Lim LZ, Song J. Dynamic principle for designing antagonistic/agonistic molecules for EphA4 receptor, the only known ALS modifier. ACS Chem Biol. 2015; 10: 372–378. doi: 10.1021/cb500413n PMID: 25334011

32. Akhvlediani T, Kvirkvelia N, Shakarishvili R, Tsertsvadze T. Als-like syndrome in the patient with chronic hepatitis C. Georgian Med News. 2009; (172–173:): 70–72. PMID: 19644195

33. Goyal S, Gupta G, Qin H, Upadya MH, Tan YJ, Chow VT, et al. VAPC, an human endogenous inhibitor for hepatitis C virus (HCV) infection, is intrinsically unstructured but forms a "fuzzy complex" with HCV NS5B. PLoS One. 2012; 7 7: e40341. doi: 10.1371/journal.pone.0040341 PMID: 22815741

34. Dutta S, Gupta G, Choi YW, Kotaka M, Fielding BC, Song J, et al. The variable N-terminal region of DDX5 contains structural elements and auto-inhibits its interaction with NS5B of hepatitis C virus. Biochem J. 2012; 446: 37–46. doi: 10.1042/BJ20120001 PMID: 22640416

35. Qin H, Shi J, Noberini R, Pasquale EB, Song J. Crystal structure and NMR binding reveal that two small molecule antagonists target the high affinity ephrin-binding channel of the EphA4 receptor. J Biol Chem. 2008; 283: 29473–29484. doi: 10.1074/jbc.M804114200 PMID: 18708347

36. Qin H, Noberini R, Huan X, Shi J, Pasquale EB, Song J. Structural characterization of the EphA4-Ephrin-B2 complex reveals new features enabling Eph-ephrin binding promiscuity. J Biol Chem. 2010; 285: 644–654. doi: 10.1074/jbc.M109.064824 PMID: 19875447

37. Petty A, Myshkin E, Qin H, Guo H, Miao H, Tochtrop GP, et al. A small molecule agonist of EphA2 receptor tyrosine kinase inhibits tumor cell migration in vitro and prostate cancer metastasis in vivo. PLoS One. 2012; 7: e42120. doi: 10.1371/journal.pone.0042120 PMID: 22916121

38. Huan X, Shi J, Lim L, Mitra S, Zhu W, Qin H, et al. Unique structure and dynamics of the EphA5 ligand binding domain mediate its binding specificity as revealed by X-ray crystallography, NMR and MD simulations. PLoS One. 2013; 8: e74040. doi: 10.1371/journal.pone.0074040 PMID: 24086308

39. Tao YJ, Ye Q. RNA virus replication complexes. PLoS Pathog. 2010; 6: e1000943. doi: 10.1371/journal.ppat.1000943 PMID: 20661480

40. Denison MR. Seeking membranes: positive-strand RNA virus replication complexes. PLoS Biol. 2008; 6: e270. doi: 10.1371/journal.pbio.0060270 PMID: 18959488

41. Bartenschlager R, Lohmann V, Penin F. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. Nat Rev Microbiol. 2013; 11: 482–496. doi: 10.1038/nrmicro3046 PMID: 23748342

42. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol. 1999; 293: 321–331. PMID: 10550212

43. Song J, Chen Z, Xu P, Gingras R, Ng A, Leberer E, et al. Molecular interactions of the Gb binding domain of the Ste20p/PAK family of protein kinases. An isolated but fully functional Gb binding domain from Ste20p is only partially folded as shown by heteronuclear NMR spectroscopy. J Biol Chem. 2001; 276: 41205–41212. PMID: 11509560

44. Li M, Song J. The N- and C-termini of the human Nogo molecules are intrinsically unstructured: bioinformatics, CD, NMR characterization, and functional implications. Proteins. 2007; 68: 100–108. PMID: 17397058

45. Liu J, Song J. A novel nucleolar transcriptional activator ApLLP for longterm memory formation is intrinsically unstructured but functionally active. Biochem Biophys Res Commun. 2008; 366: 585–591. PMID: 18078811

46. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem Sci. 2008; 33: 2–8. PMID: 18054235

47. Liu J, Li M, Ran X, Fan JS, Song J. Structural insight into the binding diversity between the human Nck2 SH3 domains and proline-rich proteins. Biochemistry. 2006; 45: 7171–7184. PMID: 16752908

48. Wang W, Lim L, Baskaran Y, Manser E, Song J. NMR binding and crystal structure reveal that intrinsically-unstructured regulatory domain auto-inhibits PAK4 by a mechanism different from that of PAK1.

Biochem Biophys Res Commun. 2013; 438: 169–174. doi: 10.1016/j.bbrc.2013.07.047 PMID: 23876315

49. Adachi T, Ago H, Habuka N, Okuda K, Komatsu M, Ikeda S, et al. The essential role of C-terminal residues in regulating the activity of hepatitis C virus RNA-dependent RNA polymerase. Biochim Biophys Acta. 2002; 1601: 38–48. PMID: 12429501

50. Carroll SS, Sardana V, Yang Z, Jacobs AR, Mizenko C, Hall D, et al. Only a small fraction of purified hepatitis C RNA-dependent RNA polymerase is catalytically competent: implications for viral replication and in vitro assays. Biochemistry. 2000; 39: 8243–8249. PMID: 10889032

51. Ranjith-Kumar CT, Gutshall L, Sarisky RT, Kao CC. Multiple interactions within the hepatitis C virus RNA polymerase repress primer-dependent RNA synthesis. J Mol Biol. 2003; 330: 675–685. PMID: 12850139

52. Nishimura AL, Al-Chalabi A, Zatz M. A common founder for amyotrophic lateral sclerosis type 8 (ALS8) in the Brazilian population. Hum Genet. 2005; 118: 499–500. PMID: 16187141

53. Chen HJ, Anagnostou G, Chai A, Withers J, Morris A, Adhikaree J, et al. Characterization of the properties of a novel mutation in VAPB in familial amyotrophic lateral sclerosis. J Biol Chem. 2010; 285: 40266–40281. doi: 10.1074/jbc.M110.161398 PMID: 20940299

54. Bastos AF, Orsini M, Machado D, Mello MP, Nader S, Silva JG, et al. Amyotrophic lateral sclerosis: one or multiple causes? Neurol Int. 2011; 3: e4. doi: 10.4081/ni.2011.e4 PMID: 21785676