HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity

Grégoire de Bisschop, Melissa Ameur, Nathalie Ulryck, Fatima Benattia, Luc Ponchon, Bruno Sargueil, Nathalie Chamond

HIV-1 gRNA, a biological substrate, uncovers the potency of DDX3X biochemical activity

Grégoire de Bisschop[1], Mélissa Ameur[1], Nathalie Ulryck [2], Fatima Benattia[1], Luc Ponchon[1], Bruno Sargueil[3] and Nathalie Chamond [3]

[1] CNRS UMR8015, Université Paris Descartes, Paris Cedex 06, 75270, France
[2] Present address: nathalie.ulryck@polytechnique.edu

[3] To whom correspondence should be addressed. Tel: 33 1 70 64 94 06; Fax: 33 1 53 73 99 25; Email: nathalie.chamond@parisdescartes.fr or bruno.sargueil@parisdescartes.fr

Highlights

Complete characterization of DDX3X ATPase activity with a biological substrate

The binding capacity of DDX3X and its enzymatic activity are dissociated properties

A specific fragment of HIV-1 gRNA stimulates DDX3X activity

The N-and C-terminal regions of DDX3X are involved in the specific binding of HIV-1 gRNA

ABSTRACT

DEAD-box helicases are key players all along the existence of many RNAs and ribonucleoproteins assisting their synthesis, folding, function and even their degradation or disassembly. They have been implicated in various phenomena, and it is often difficult to rationalize their molecular roles from *in vivo* studies. Once purified *in vitro,* most of them only exhibit a marginal activity and poor specificity. The current model is that they gain specificity and activity through interaction of their intrinsically disordered domains with specific RNA or proteins. DDX3X is a DEAD-box cellular helicase that has been involved in several steps of the HIV viral cycle, including transcription, RNA export to the cytoplasm and translation. In this study, we investigated DDX3X biochemical properties in the context of a biological substrate. DDX3X was overexpressed, purified and its enzymatic activities as well as its RNA binding properties were characterized using both model substrates and a biological substrate, HIV-1 gRNA. Biochemical characterization of DDX3X in the context of a biological substrate identifies HIV-1 gRNA as a rare example of specific substrate and unravels the extent of DDX3X ATPase activity. Analysis of DDX3X binding capacity indicates an unexpected dissociation between its binding capacity and its biochemical activity. We further demonstrate that interaction of DDX3X with HIV-1 RNA relies both on specific RNA determinants and on the disordered N- and C-terminal regions of the protein. These findings shed a new light regarding the potentiality of DDX3X biochemical activity supporting its multiple cellular functions.

Keywords: DDX3X, HIV-1, ATPase, biological substrate, RNA specificity

INTRODUCTION

RNA helicases are involved in all major steps of RNA metabolism from transcription, pre-mRNA processing to RNA transport and decay. Helicases are classified into super-families and families according to conserved sequence motifs, domain composition and oligomeric state [1]. The DEAD-box protein family, which belongs to superfamily 2, forms the largest group of helicases [2]. DEAD-box proteins are characterized by the signature sequence D-E-A-D (Asp-Glu-Ala-Asp) in their conserved Walker B motif [3] and are present in numerous prokaryotes and in all eukaryotes [4,5]. DEAD-box proteins are ATP-dependent RNA binding proteins that remodel RNA structures and RNA-protein complexes [6,7]. DDX3 (encoded by the *ddx3x* gene) and Ded1p, its yeast orthologue, define the Ded1 / DDX3 subfamily with additional conserved N- and C-terminal regions involved in RNA and protein interactions [8].

DDX3 is involved in numerous physiological cellular processes such as cell cycle progression, apoptosis, RNA metabolism, from transcription to translation, splicing, nuclear export and mRNA degradation [8–10]. In addition, DDX3 has been identified as a key molecule in pathological situations such as in various cancer types [11] as well as during diverse viral infections [12,13], including HIV-1 [14]. DDX3 is essential to HIV-1 replication through its role both in the CRM1-mediated nuclear export of the unspliced HIV-1 RNA [14,15] and in the viral genome translation [16]. Indeed, DDX3 shuttles from the nucleus to the cytoplasm where it allows the ATP-dependent formation of cytoplasmic granules that are required for HIV-1 translation [17].

HIV-1 full length RNA, known as the genomic RNA (gRNA), is multifunctional. It acts both as an mRNA that is translated to yield the Gag and Gag-Pol polyproteins and as a genome that is encapsidated within the virions. Translation of eukaryotic mRNA relies on the recognition of the 5'-cap structure to recruit the 43S complex (composed of the 40S subunit, tRNA, eIF2 and eIF3). This is mediated by the eIF4F complex that allows cap recognition through eIF4E and unwinding of mRNA secondary structures by the eIF4A helicase. The eIF4G platform coordinates 40S and mRNA association through its connection with eIF4E, eIF3 and the Poly(A) Binding Protein (PABP). The 5'-UTR of HIV-1 gRNA is long, highly structured and thus predicted to impede the scanning step of the canonical cap-dependent initiation pathway. More specifically, the presence of the 5' proximal extra-stable stem loop TAR strongly inhibits translation of HIV-1 gRNA [18]. The exact molecular mechanism responsible for this inhibition remains undetermined; it could hinder the cap accessibility to eIF4E, and/or prevent ribosome binding or progression on the mRNA. Nonetheless, HIV-1 can initiate translation either by internal recruitment of the ribosome through RNA domains called IRESs (internal ribosome entry sites) or by the classical cap-dependent mechanism [19]. DDX3 has been shown to be required for HIV-1 gRNA efficient translation *in cellulo*. More precisely, the helicase is required for 43S initiation complexes recruitment to the 5' cap structure of HIV-1 gRNA [16]. However, its exact role in HIV-1 translation and in the initiation mechanisms involved are yet to be defined.

DEAD-box helicases share 12 conserved motifs that are involved in ATP binding and hydrolysis, in RNA binding and in the coupling of ATP- and RNA-binding; which are biochemical properties supporting their biological functions. Ded1p displays ATP-

dependent helicase and RNA-dependent ATPase activities [20] as well as additional properties including RNP remodeling [21,22], stable RNA binding [23], AMP sensing [24] and strand annealing [25] activities. The accepted model of RNA duplex unwinding is by "local strand separation" [26] where DEAD-box helicases load directly onto the duplex and separate a limited number of base pairs in an ATP-dependent fashion (ATP binding). ATP hydrolysis is necessary for RNA release and enzyme recycling. More recently, it has been shown that Ded1p activity relies on the assembly of three Ded1p “protomers” [27]. Two “protomers” loaded onto single-stranded RNA stimulate the binding of a third one onto duplex RNA, allowing for strand separation.

The proof of concept of DDX3 basic properties (RNA-dependent ATPase and RNA unwinding activities) was first attained by Yedavalli et al. in 2004 [14] although the protein had been identified in 1998 by Park et al [28]. Nonetheless, through three seminal papers approaching the biochemical characterization of DDX3 [14,29,30], the recombinant protein was described as displaying distinct features compared to its yeast orthologue. For instance, DDX3 was reported to present a relatively high level of RNA-independent ATPase activity and an ATPase activity that could be stimulated by RNA and DNA (DNA-stimulated ATPase activity is unusual for DEAD-box helicases). In addition, a DDX3 fragment lacking critical RNA binding motifs and motif VI (which contains the catalytically critical arginine finger) maintained intrinsic ATPase activity. Finally, DDX3 was reported to hydrolyze all NTPs and not to be restricted to ATP. More recently, Sharma et al. performed a thorough characterization of DDX3 helicase activity using model substrates, thereby identifying DDX3 as a "true" DEAD-box helicase although the enzyme displays some specific features as compared to its yeast

orthologue Ded1p (reduced strand annealing activity and preference for 3' single-strand RNA) [31].

To appreciate the link between the biochemical properties of DDX3 and its role in HIV-1 gRNA translation, we explored the relevance of DDX3 biochemical parameters in the context of a biological substrate. We produced the recombinant protein and performed a thorough characterization of its ATPase activity when stimulated by HIV-1 gRNA; activity of which cannot be fully attained with model substrates. The production of the DQAD mutant allowed us to analyze the binding properties of DDX3 to RNA and to unravel an unexpected dissociation between its binding capacity and its biochemical activities. We further demonstrate that interaction of DDX3X with HIV-1 RNA relies both on specific RNA determinants and on the disordered N- and C-terminal regions of the protein. Altogether, our results specify the biochemical properties of DDX3X and the RNA requirements for a biological target.

MATERIAL AND METHODS

Cloning, expression and purification

The pET21 construct containing the complete open reading frame of hDDX3X was a kind gift of T. Olhmann and R. Soto-Rifo. The coding sequence was sub-cloned into pETM41 expression vector leading to an N-terminally hexahistidine-tagged fusion protein of MBP (Maltose Binding Protein) and DDX3, with a TEV protease recognition site in between. QuickChange site-directed mutagenesis (Stratagene) was used to generate the ATPase-deficient Motif II mutant (DQAD). Wild-type and mutant proteins were produced following the same protocol. Proteins were expressed in BL21 Star

(DE3) cells (Invitrogen), grown in LB medium and protein expression was induced at OD_{600} = 0.6 upon addition of 0.1 mM IPTG. Bacteria were grown for 2 hours at 37°C. Cells were harvested by centrifugation, cell pellets were resuspended in Buffer A containing 50 mM Tris-Cl pH 8 / 300 mM KCl / 10 mM Imidazole in the presence of proteases inhibitor (Roche) and cell disruption was obtained after 3 passages in the Eaton-press. The lysate was cleared at 24000 rpm at 10°C for 20 min and the supernatant was loaded on a Ni-NTA His Trap column (GE Healthcare), pre-equilibrated in Buffer A. The column was washed with 10 CV in Buffer A and 10 CV in Buffer A containing 1 M KCl to avoid nucleic acids contaminations. Recombinant protein was eluted in a gradient of Buffer A supplemented with 500 mM Imidazole. Fractions containing MBP-DDX3 were pooled, dialyzed overnight against 20 mM Tris-Cl pH 8 / 200 mM KCl / 1 mM DTT / 5 % glycerol (Buffer B) and loaded on a MonoQ 10/100 GL column (GE Healthcare) pre-equilibrated in Buffer B. The protein was eluted in a linear gradient to MonoQ buffer supplemented with 1 M KCl. Whenever indicated, the protein was cleaved with the TEV protease to remove the MBP fusion protein. The protein was further purified by the use of an Heparin column pre-equilibrated in Buffer C (20 mM Tris-Cl pH8, 200 mM KCl, 5 % glycerol and 1 mM DTT) and a linear gradient to Buffer C supplemented with 2 M KCl. Fractions containing pure protein were pooled and dialyzed against 20 mM Tris-Cl pH8, 250 mM KCl, 1 mM DTT and 30 % glycerol. All purifications steps were monitored by 8 % SDS-PAGE and Coomassie-blue staining. The purified protein was quantified by performing a 230 - 320 nm OD spectrum and examination of the OD 260/280. Purified Ded1p is a kind gift of K. Tanner and J. Banroques.

ATP hydrolysis assay

ATPase assays were performed in 20 µl reactions containing 100-500 nM of purified protein in 20 mM Tris-Cl pH 8, 80 mM KCl, 2.5 mM $MgCl_2$ and 1 mM DTT in the presence of 1-1000 nM RNA. The reaction was initiated by the addition of an ATP solution (1-3200 µM of cold ATP/$MgCl_2$ and 3 nM of 800 Ci / mmol α^{32}P-ATP as the final concentrations in the reaction) and allowed to proceed for 10-20 minutes at 37°C. Aliquots (2 µl) were removed at various time points, hydrolysis products were separated by thin layer chromatography on polyethylenimine cellulose plates in 0.5 M KH_2PO_4 pH 3.4, exposed and analyzed using a BAS-5000 scanner (Fujifilm). For competition assays, the ATP solution (10 µM) was supplemented by 100 µM of the different NTPs. The different nucleic acids used for ATPase activity stimulation are indicated in the Supplemental Table 1. All RNAs were directly transcribed using the T7 RNA polymerase from polymerase chain reaction (PCR) products containing the T7 RNA polymerase promoter sequence (5'-TAATACGACTCACTATAG-3') and purified by size exclusion chromatography. The apparent affinity for the ATP substrate and the apparent rate of ATP hydrolysis were calculated by measuring the variation of initial velocities of the ATPase reaction, as a function of ATP concentration. Data were analyzed according to the Michaelis-Menten equation: $V = (k.E_0.S^n)/(Km + S^n)$, where k is the catalytic rate, E_0 is the input enzyme concentration, S is the ATP concentration and n is the cooperativity index. Experimental data and statistical analyses were performed using GraphPad Prism.

Duplex unwinding assay

The RNA:RNA and RNA:DNA substrates for the unwinding assays were formed by annealing a 5'-end radiolabeled short RNA or DNA oligonucleotide to either the 5' or 3' end of a longer RNA oligonucleotide (Supplemental Table 1). Oligonucleotides were purchased from Eurogentec. Unwinding assays were performed in unwinding buffer (40 mM Tris-Cl pH8, 50 mM NaCl, 0.5 mM $MgCl_2$, 2 mM DTT and 0.01 % NP40) in 20 µl reactions containing 100 nM of purified protein, 2 mM ATP/$MgCl_2$ and 5 nM of the different substrates (as indicated in the Figures). Reactions were initiated by the addition of ATP, incubated at 37°C and stopped with a solution containing 1 % SDS, 50 mM EDTA, 0.1 % xylene cyanol, 0.1 % Bromophenol Blue and 20 % glycerol. Samples were loaded on 15 % native acrylamide gels, run at 4°C for 1 h, the gels were dried, exposed and analyzed using a BAS-5000 scanner and the software MutliGauge.

In vitro transcription, mobility shift assays and filter binding assays, were essentially performed as in [32,33]. These methods are detailed in the Supplemental Material and Methods Section.

RESULTS

Production of soluble and active DDX3

To facilitate the biochemical characterization of DDX3X, we overexpressed the protein in *E. coli*. In first attempts, the full-length protein was cloned in pET vectors fused to a His_6-tag in C- or N-terminus. Using standard protocols for protein expression and extraction, we collected large amounts of DDX3X but entirely in the insoluble fraction. All our efforts aiming at either reducing protein expression or synthesis rate, or facilitating

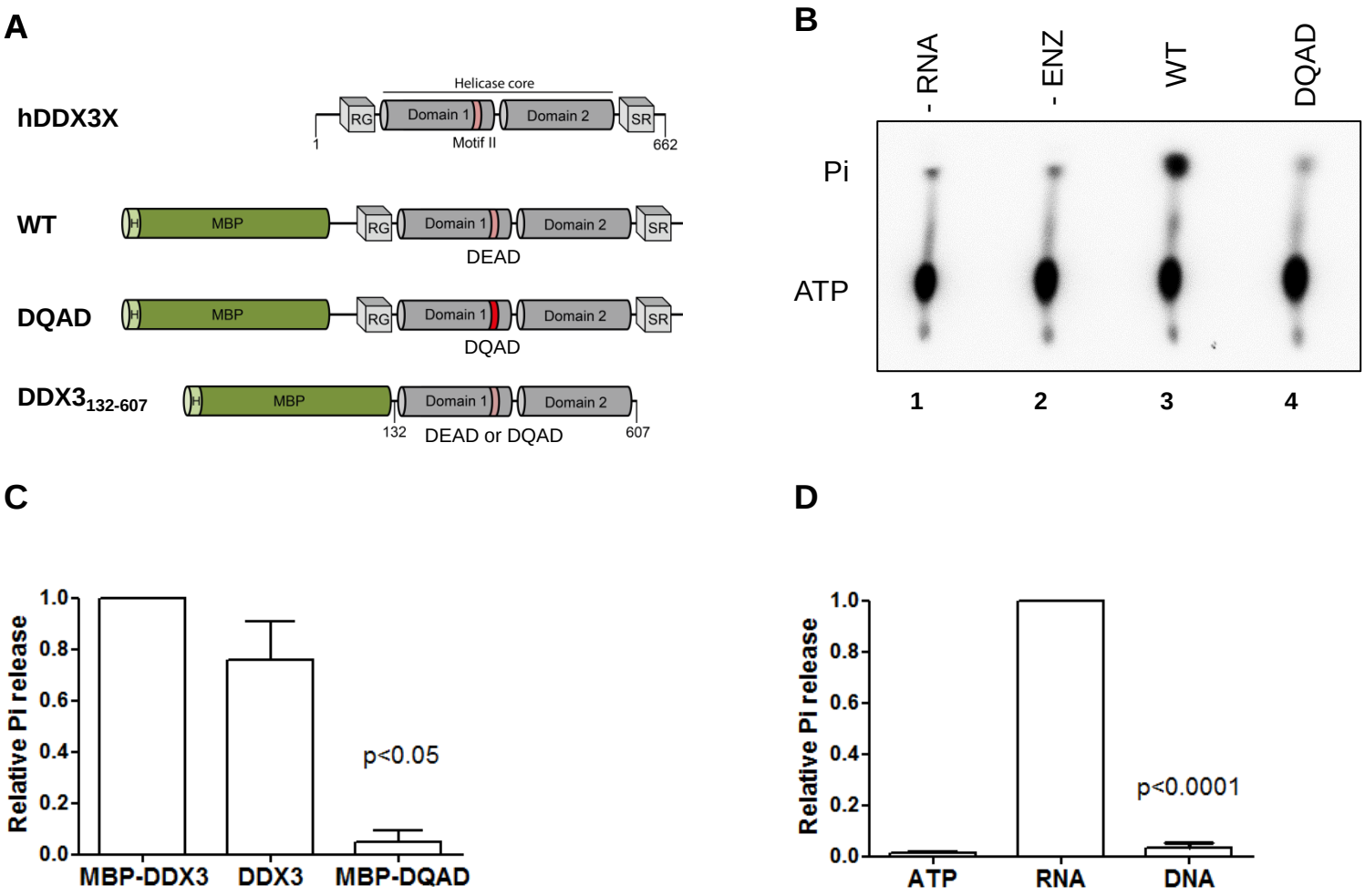


FIGURE 1: Purification of an active recombinant DDX3 protein. (A) Schematic representation of hDDX3X and of the derivative constructs used in this study. H, His6x tag; RG, arginine and glycine rich N-terminal domain ; SR, serine rich C-terminal domain. (B) The ATPase activity of wild-type (WT) or mutant (DQAD) MBP-DDX3 was assessed by thin layer chromatography measurement of Pi release in the absence (lane 1) or presence (lanes 2-4) of HIV-1 RNA. (C) Quantitation of the ATPase activity of DDX3 and its variants. (D) Quantitation of the Pi released by 100 nM protein incubated in the presence of 100 nM HIV-1 RNA or its corresponding dsDNA and 100 µM ATP. The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from a paired t-test. * p<0.05.

protein folding by chaperonin (GroEL/ES) overexpression failed to yield the protein in the soluble fraction. Similarly, varying the lysis conditions or solubilizing the protein with denaturants such as urea or guanidium chloride did not result in the production of sufficient amounts of soluble active recombinant protein. We then utilized the His_6-Maltose Binding Protein (MBP) fused with the N-terminus of DDX3X (Figure 1A). Most of the protein was in the soluble fraction of bacterial extracts. The recombinant protein was purified on a nickel-agarose column, its apparent molecular mass (about 120 kD) was in agreement with its estimated weight and the identity of the protein was confirmed by western blot analysis using anti-DDX3X antibody (data not shown). Fractions enriched in MBP-DDX3X were further pooled and applied onto a MonoQ column. The resulting recombinant protein was about 90% pure as judged by Coomassie staining (Supplemental Figure S1). The absence of nucleic acids contamination was obtained by performing a wash step with 1M salt during the nickel column purification followed by the MonoQ column which is known to trap nucleic acids; and ascertained by a 230-320 nm OD spectrum (see Methods). To exclude a drastic influence of the MBP-tag on DDX3X properties, we performed an additional TEV-cleavage followed by a third purification step on a heparin column to separate the cleaved MBP-tag and the TEV protease from DDX3X. In addition, to corroborate the importance of the DEAD motif on DDX3X properties, we constructed a plasmid harboring the DQAD motif. Molecular replacement of the glutamate moiety by a glutamine in Motif II results in the absence of ATP hydrolysis but does excessively interfere neither with ATP binding nor with nucleic acid interaction [34]. The DQAD mutant was purified in the same conditions as the wild-type protein. The yield and purity were similar to those obtained with the wild-type (Supplemental Figure S2).

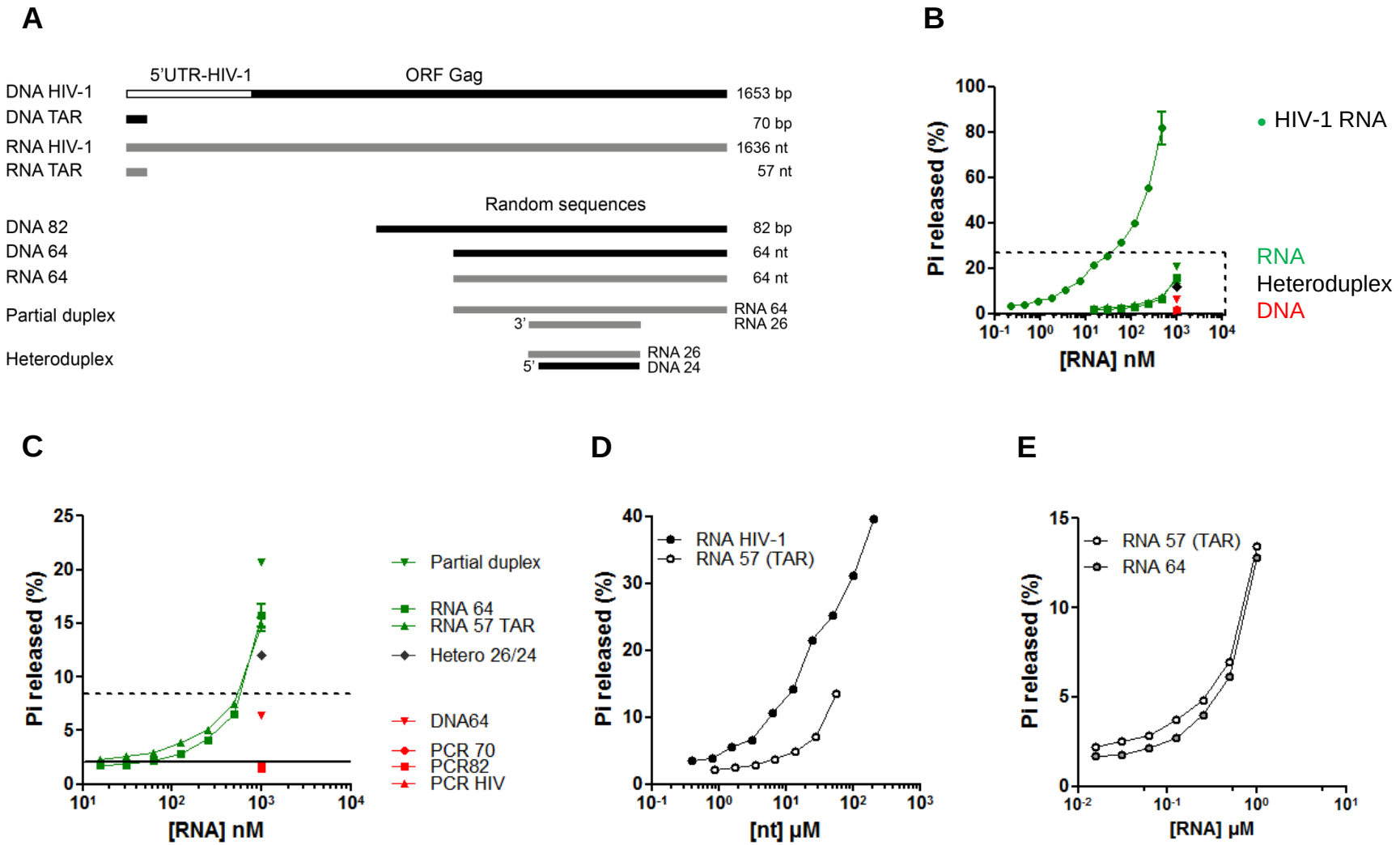


FIGURE 2: Stimulation of DDX3 ATPase activity with various RNA or DNA substrates. (A) Schematic representation of the different substrates used. (B) Pi release was measured in the presence of RNA HIV-1 (green circle), less than 100 nt RNA (green), DNA (red) or heteroduplex NA substrates (black) and 200 nM ATP and 600 nM enzyme. For clarity, "small" substrates (dotted square) are enlarged in (C). (D) and (E) Pi release was measured in the presence of the indicated substrates at different concentrations expressed as nucleotide (nt) or molar (RNA) concentrations to account for the size difference whenever relevant.

DDX3 is an active RNA-dependent ATPase

The ATPase activity was measured by TLC, in which ATP hydrolysis is directly assessed by Pi release quantification. Figure 1B represents typical data generated by the ATPase assay. In absence of RNA or of the enzyme, only traces of Pi were detected contrary to what is observed in the presence of DDX3X and RNA; therefore demonstrating that DDX3 has no intrinsic ATPase activity (Figure 1B, lane 3)which is consistent with the absence of contaminating nucleic acids in our preparation. This activity was not observed with the DDX3X-DQAD mutant (Figure 1B, lane 4) thereby demonstrating that measured ATPase activity is that of DDX3X. Importantly, when the MBP-tag was removed from DDX3X, similar yields of Pi release were obtained indicating that the presence of the MBP at the N-terminus of DDX3X does not alter its activity (Figure 1C). Since the cleaved protein was less stable in our hands, for practical reason, but also to ensure that the protein does not precipitates during the assays, all subsequent experiments were performed with MBP-DDX3X (further referred to as DDX3). In addition, when dsDNA was used to stimulate the ATPase activity, no significant increase in Pi release was observed (Figure 1D) therefore demonstrating that DDX3 ATPase activity is strictly RNA-dependent.

To test for a possible specificity of the ATPase reaction toward the nucleic acid substrate, we compared a variety of nucleic acids (Figure 2A, sequences are indicated in the Supplemental Table 1) that differ by their nature (DNA or RNA), length or structure (ss, single-stranded or ds, double-stranded). As shown in Figures 2B and 2C, while DNAs are unable to stimulate DDX3 ATPase activity, all the RNAs tested activated the ATPase activity, although with different levels (ranging from 15 to 90% of Pi release).

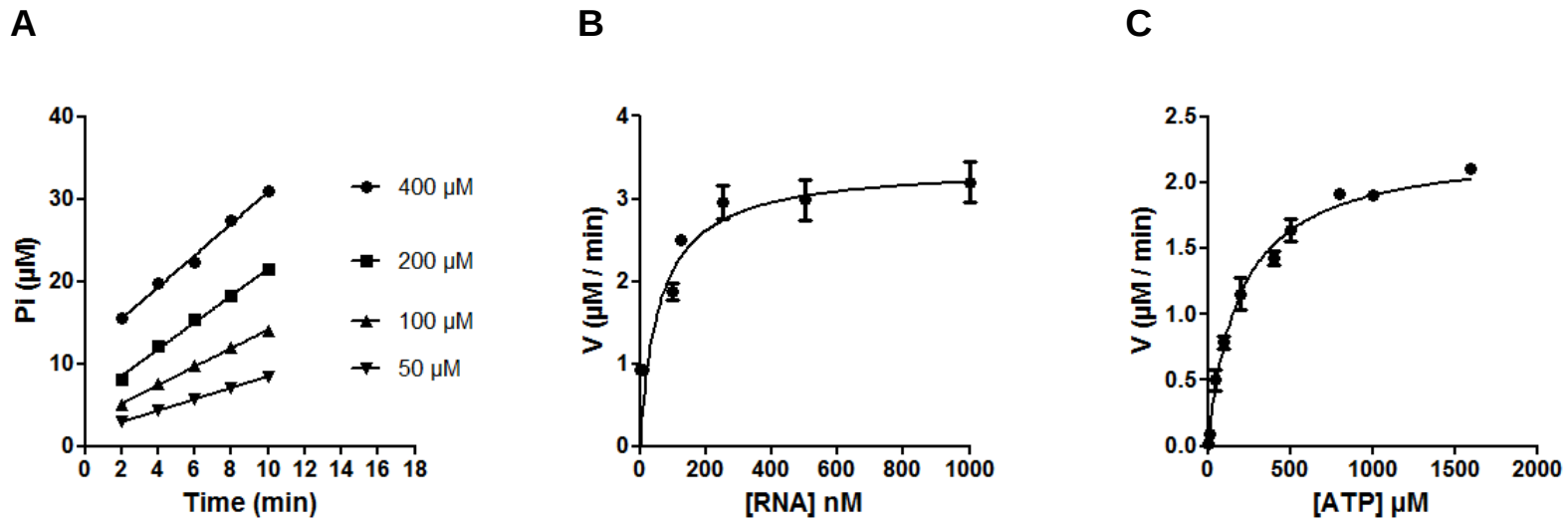


FIGURE 3: Determination of the kinetic parameters of DDX3 ATPase activity. ATPase reactions were performed in the presence of 100 nM DDX3, 100 nM RNA HIV-1 and 50-400 µM ATP. (A) The linear regression of Pi release versus time allows for the initial rate (V) determination at different ATP concentrations. (B) RNA stimulation of DDX3 ATPase activity was measured for different concentrations of RNA (10-1000 nM) with saturating ATP (1-2 mM). (C) DDX3 exhibits Michaelis-Menten kinetics with an apparent K_{mATP} = 196 ± 26 µM and an apparent V_{max} of 2.3 ± 0.1 µM.min^{-1}. The values represent the average and standard deviation of at least three independent experiments.

The highest activity was observed with RNA HIV-1, which could be the result of the size difference of the RNAs used or of their structural status, or of both. In order to evaluate further if one of these parameters influences the ATPase activity of DDX3, we compared the RNA stimulation obtained with RNA HIV-1 and RNA TAR (57 nt long) when the amount of RNA is expressed as the nucleotide concentration. As can be observed in Figure 2D, for equivalent nucleotides concentrations, RNA HIV-1 induces the release of twice the amount of Pi as compared to RNA TAR (25 % and 13 %, respectively). Importantly, RNA HIV-1 contains the extra-stable stem loop TAR that has been shown to be a critical determinant for HIV-1 gRNA interaction with DDX3 [17]. However, in our ATPase assay, the isolated TAR is not sufficient to recapitulate HIV-1 gRNA stimulation of DDX3 activity. In addition, when we compared two RNAs that are similar in size (RNA TAR and RNA 64) but that differ from their structural status (extra-stable TAR stem loop vs essentially single-stranded RNA of random sequence), both species stimulated the ATPase activity of DDX3 to a similar extent. Taken together, our results suggest that single- or double-stranded RNA is not an optimal substrate to stimulate DDX3 ATPase activity but rather, more complex RNAs alternating single- and double-stranded regions.

Next, we performed Michaelis-Menten kinetics and characterized the ATP hydrolysis performed by DDX3. DDX3 (100 nM) in the presence of HIV-1 RNA was incubated with different concentrations of ATP-γ^{32}P (Supplemental Figure S3). The linear regression of Pi release versus time allows for the initial rate determination (V) at different ATP concentrations (Figure 3A). First, kinetics were assayed in the presence of various RNA concentrations (1 - 1000 nM) and saturating ATP (1 - 2 mM ATP) conditions. As can be observed in Figure 3B, saturation of the apparent Vmax is attained

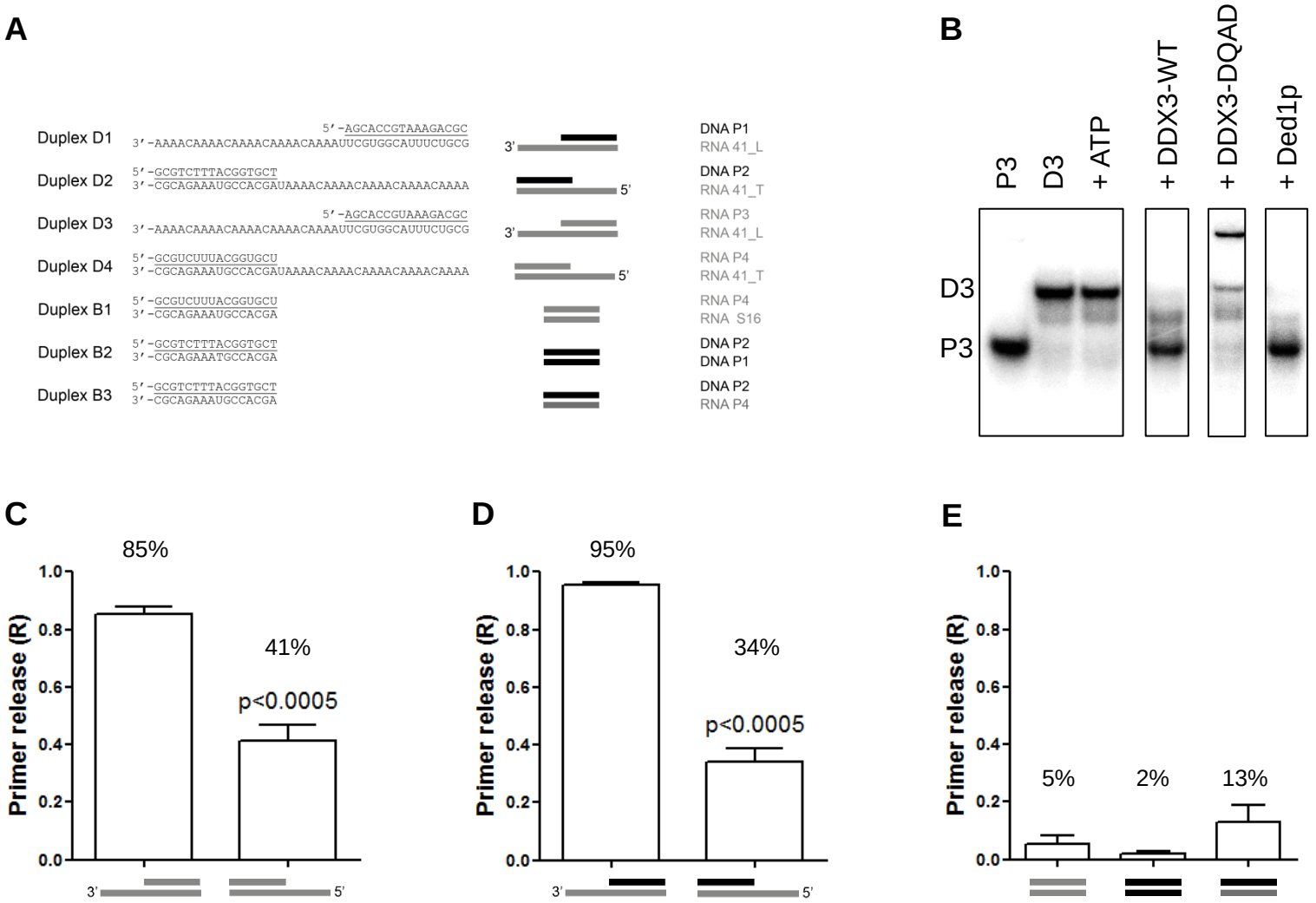


FIGURE 4: DDX3 is an active helicase. (A) Sequences and schematic representation of the substrates used, from [27]. (B) Representative PAGE (complete gels in Supplemental Figure S4) for unwinding reactions performed in the absence (+ATP) or in the presence of DDX3 (+DDX3-WT), of the DQAD mutant (+DDX3-DQAD) or of Ded1p (+Ded1p). (C) (D) (E) Quantification of primer release by 100 nM of DDX3 in the presence of 5 nM substrate and saturating 2 mM ATP. The cartoons represent the helicase substrates described in (A). The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired t-test.

in the presence of HIV-1 RNA with a $k_{1/2\ [RNA]}$ of 59.6 ± 21.0 nM. This result differ by several orders of magnitude from what was recently reported using model substrates ($K_{1/2[RNA]}$ of 2.3 to 3.4 µM, Sharma et al. 2017). It is important to note here that we and others could not attain RNA saturating conditions with model substrates, even at high concentrations (up to 10 µM). We next tested ATP hydrolysis at unsaturated to saturated ATP concentrations to determine the kinetic parameters of DDX3 towards ATP. As can be observed in Figure 3C, DDX3 exhibits Michaelis-Menten kinetics with an apparent K_{mATP} of 193 ± 26 µM and an apparent K_{cat} of 23 ± 10 min^{-1}.

DDX3 is an active ATP-dependent RNA helicase

To validate the helicase activity of our recombinant protein, duplexes commonly used (Figure 4A, [27]), were incubated in the presence of saturating ATP (2 mM) and 140 nM enzyme (Figure 4 and Supplemental Figure S4). As can be observed in Figure 4B, DDX3 efficiently unwinds RNA/RNA duplexes (D3) in the presence of ATP, similarly to Ded1p. The DDX3-DQAD mutant which cannot hydrolyze ATP did not display any helicase activity and this was observed independently of the substrate used. Nonetheless, we observed with the mutant a higher band shift that could reflect its interaction with the duplexes in the presence of ATP (see further). Importantly, when duplexes (RNA/RNA or RNA/DNA) with either 5' or 3' single stranded overhangs were tested, DDX3 displays a marked preference for 3'-tails, with an apparent two-fold decrease in DDX3 helicase activity whenever 5'-tailed substrates were used. Finally, blunt duplexes, independently of their nature, are not substrates for the helicase activity even if incubated for longer periods of time (data not shown). We did not evaluate further the kinetic parameters of DDX3 helicase activity. Nonetheless, taken together, these

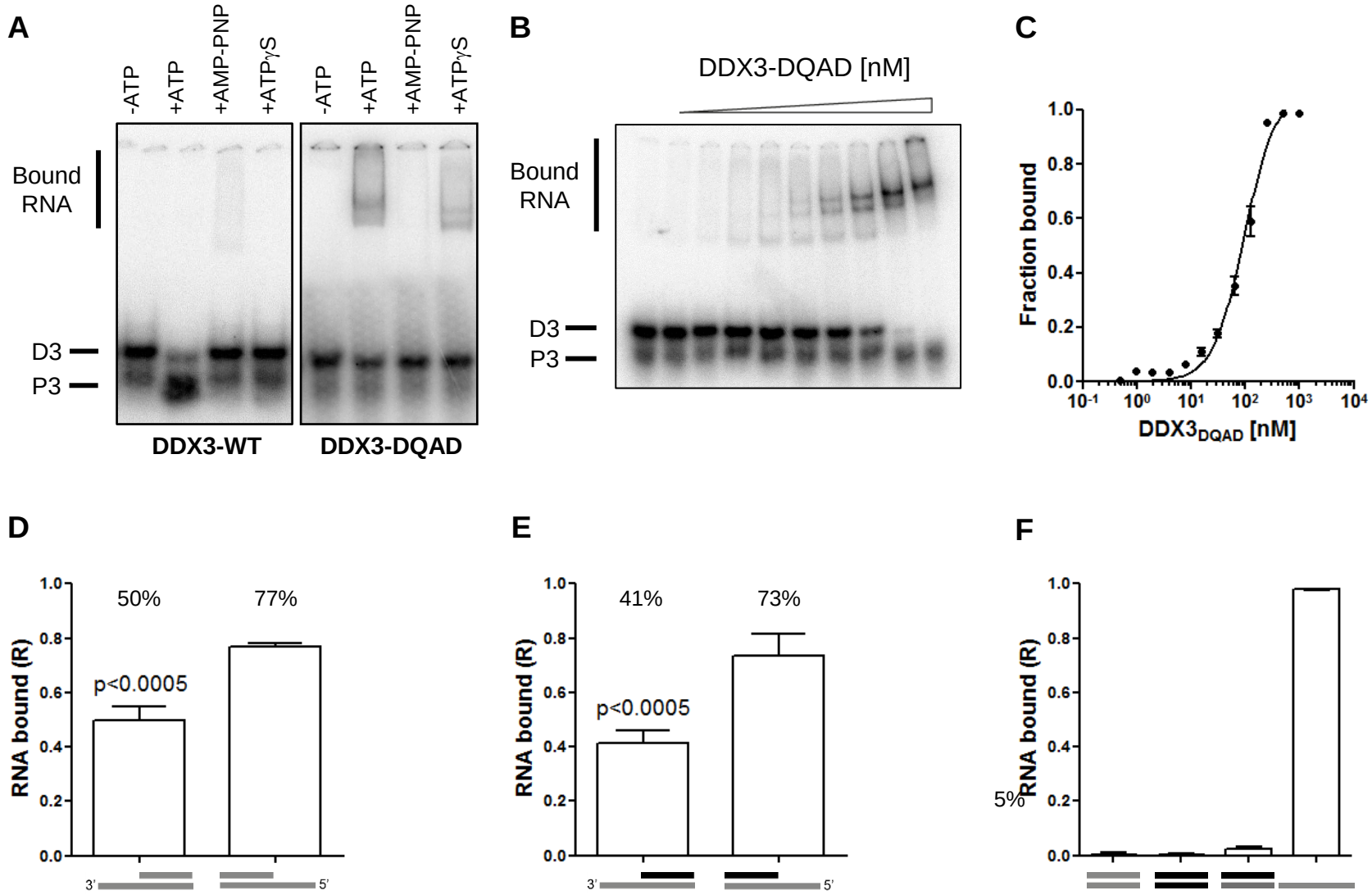


FIGURE 5: DDX3-DQAD interaction with model substrates. 16 bp RNA/RNA duplexes (D3, 5 nM) were incubated in the presence of (A) 50 nM protein or of (B) increasing concentrations of DDX3-DQAD (0 - 500 nM) and 2 mM ATP. (C) The quantitation of bound duplexes allows for the determination of the following binding parameters: Bmax = 1.00 ± 0.02, h = 1.7 ± 0.2 and Kd = 80.6 ± 5.1 nM. The different duplexes (D), (E) and (F) were tested for their capacity to form stable binary complexes. The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired t-test.

results confirm that DDX3 is an active ATP-dependent RNA helicase although it can also be active on RNA/DNA heteroduplexes, and it displays a preference for substrates presenting 3' single-stranded regions.

Characterization of DDX3 RNA binding activity on synthetic substrates

Our results on the helicase activity of DDX3 indicate that the functional binding for RNA depends on the identity of the molecule used. However, apparent affinities deduced from the helicase activity potentially reflect the actual binding step as well as steps of the ATP-hydrolysis reaction, or of the helicase reaction or of the Pi or substrate release. Therefore, to investigate the preference of DDX3 for 3'-single stranded substrates, we examined its binding capacity through mobility shift assays (Figure 5). In the presence of ATP, it is not possible to observe DDX3 bound to its RNA substrate, which could be a consequence of the helicase rapid turnover. More surprisingly, when the biochemical activities of DDX3 were compromised by the use of non- or poorly-hydrolysable ATP-analogues such as AMP-PNP or ATPγS, we could not observe stable complexes even in the presence of increased DDX3 concentrations (up to 1.7 µM, data not shown). Nevertheless, whenever the binding capacities of the DQAD mutant were assessed in the same conditions (5 nM substrate, 50 nM protein), a clear band shift was only observed in the presence of ATP and to a lesser extent in the presence of ATPγS, but not in the presence of AMP-PNP. These results suggest that the DQAD-mutant, and most probably the WT DDX3, forms complexes with its RNA substrate in the presence of ATP but not in the presence of non-hydrolysable analogs, potentially because of an incorrect geometry at the γ–phosphate position. Therefore, substrate binding was characterized by gel-shift assays performed with the DDX3-DQAD mutant (Figure 5B).

Worth noting, at least two shifted species can be observed on Figures 5A and B suggesting that at least two molecules of DDX3 can bind on these substrates. We used the standard model substrates, either the 16 bp duplexes with a 25 nt 5'-or 3'-tail, the blunt 16 bp duplexes or a 41 nt ss-RNA. Quantitative analysis of the fraction of bound duplexes (R) upon DDX3 titration in the presence of the D3 substrate (RNA/RNA substrate with a 3'-overhang) reveals an apparent equilibrium constant (K_d) of 80.6 ± 5.1 nM (Figure 5C). Surprisingly, when the different helicase substrates were tested, a two-fold increase in binding is observed with substrates presenting a 5'-tail as compared to those presenting a 3'-tail, independently of the nature of the top strand (RNA or DNA) (Figure 5D-E). This result suggests that the RNA selectivity exhibited by DDX3 helicase activity is not directly proportional to the level of RNA binding. Complexes can be observed on single-strand RNA, while they fail to form with blunt-end RNA or DNA/RNA duplexes (Figure 5F). Altogether, our data show that the presence of ssRNA is sufficient to promote RNA binding and that the presence of a ss-ds junction can modulate its efficiency. This observation is consistent with the RNA requirements described for Ded1p [20].

Determination of the shortest HIV-1 substrate necessary to activate DDX3 activity

As mentioned previously, the ATPase activity of DDX3 is only moderate on short synthetic substrates. Therefore, to evaluate the link between the RNA binding capacity and the ATPase activity of DDX3, we produced several RNA substrates derived from HIV-1 gRNA or from a control RNA (Figure 6A, Supplemental Table 1) and measured their apparent Kd by filter binding assay (Table 1). Of note, this technique is performed in presence of high monovalent salt concentration (300 mM) to prevent high

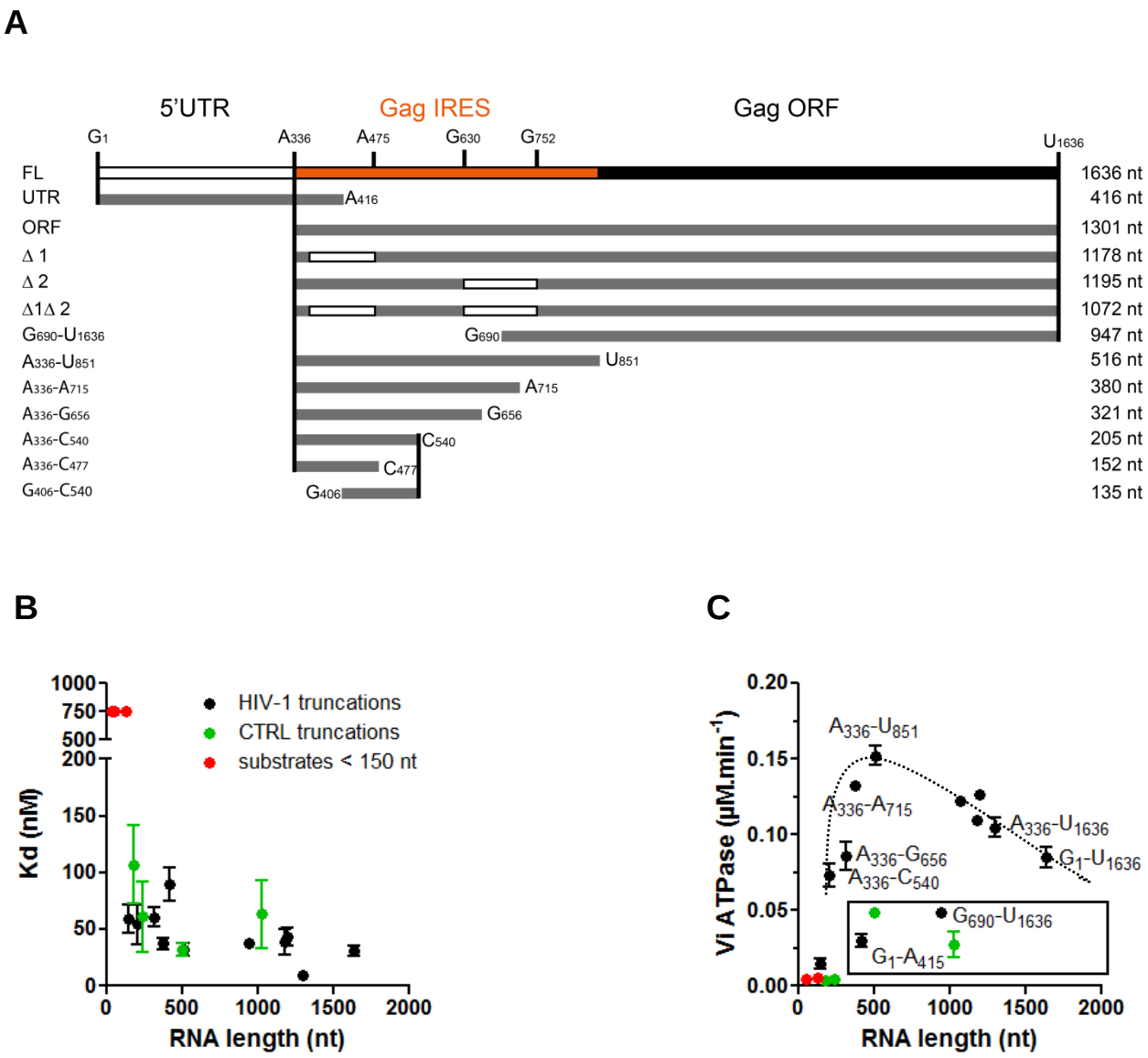


FIGURE 6: HIV-1 RNA is a preferential substrate. (A) Schematic representation of the RNA HIV-1 fragments used (on scale). (B) Representation of the results obtained in Table 1. RNAs derived from HIV-1 (black), from the luciferase control (green) or substrates < 150 nt (red) are indicated in the Figure. Whenever the Kd cannot be determined, it was artificially set at 750 nM. The results are the mean and standard deviation of at least three independent experiments. (C) Kinetic ATPase reactions were performed in the presence of 100 nM DDX3-WT, 100 nM RNA and 10 µM ATP for 10 minutes. Initial rates are indicated in Table 2 as the average and standard deviation of independent experiments.

background. Consequently, reported Kds are more elevated than those observed in low salt concentration (80 mM) using gel-shift assay, which is less applicable to long RNA substrates. The obtained Kds are not proportional to the size of the RNA molecules (Figure 6B and Table 1) but indicate that a minimal size of approximatively 200 nt is required for RNA to efficiently bind DDX3 in this assay. In addition, RNAs encoding the renilla luciferase (control RNAs) coding region display Kds that are similar to those obtained with HIV-1-derived RNAs. We then compared the capacity of the different fragments to stimulate the ATPase activity of DDX3 (Figure 6C, Table 2). As described above for short synthetic substrates HIV-1 or control RNAs of less than 200 nt fail to efficiently stimulate DDX3 ATPase activity. Over a threshold around 200 nt we do not observe any correlation between the ATPase activity and the length of the molecule used to stimulate. Interestingly, control RNAs do not stimulate the ATPase activity to the same extent as HIV-1 derived RNAs of similar size; therefore suggesting the presence of some specific elements in HIV-1. In order to identify a shorter specific substrate, HIV-1 G_1-U_{1636} was subjected to progressive deletion. Removal of the 5'-UTR in the G_1-U_{1636} (leading to A_{336}-U_{1636}) is beneficial to the ATPase activity stimulation. Worth noting, the 5'-UTR by itself is a rather poor substrate (G_1-A_{416}). Additional 3' deletions lead to an optimal fragment (A_{336}-U_{851}) corresponding to Gag-IRES (Locker et al. 2011) whereas the remaining 3' region (G_{690}-U_{1636}) fails to stimulate DDX3 ATPase activity efficiently. Further 5' or 3' deletions of this fragment drastically reduce the ATPase activity of DDX3 defining an optimal fragment of 516 nt (A_{336}-U_{851}) and a minimal fragment of 205 nt (A_{336}-C_{540}). Altogether, these results clearly demonstrate that DDX3 substrate RNA holds specific determinants.

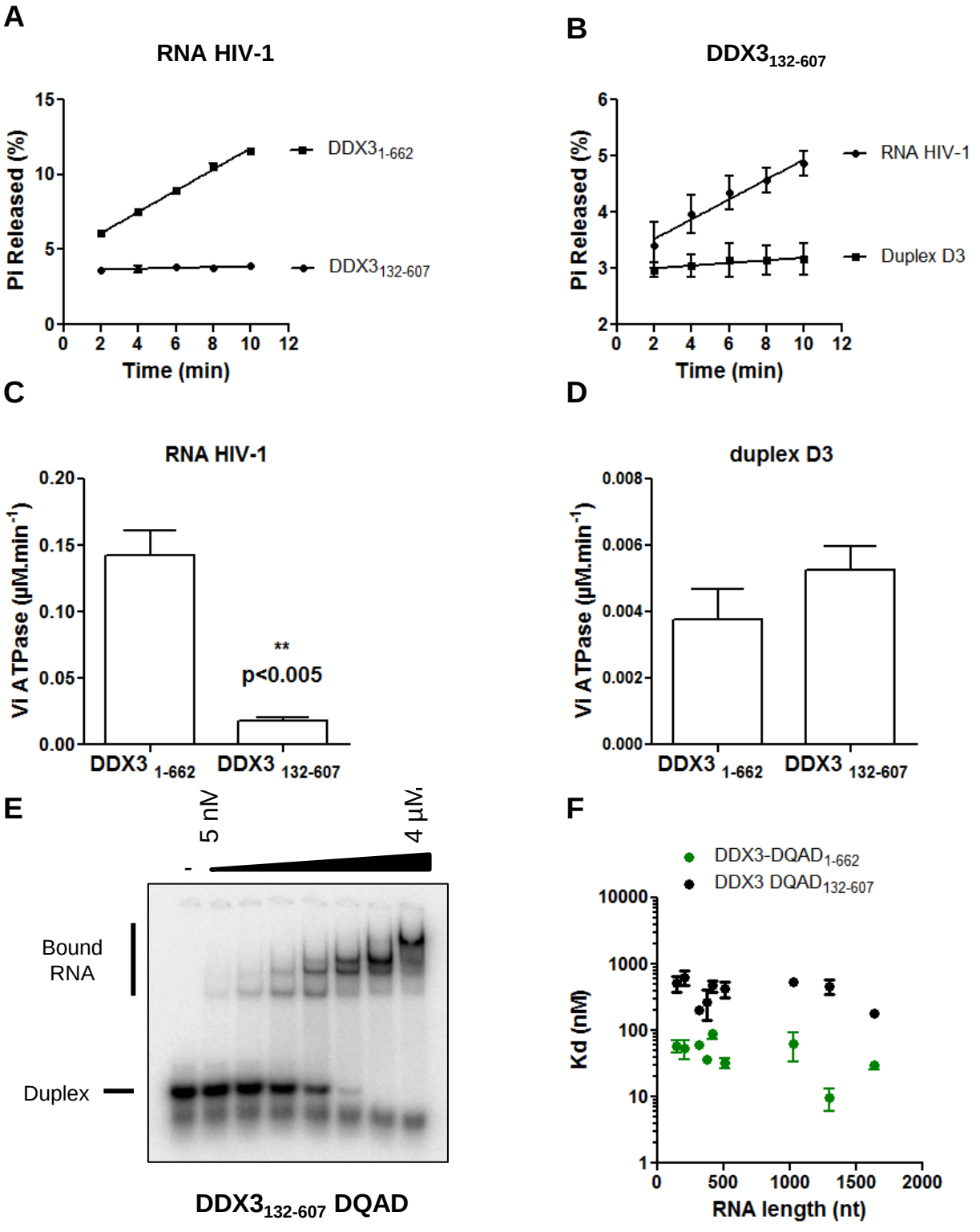


FIGURE 7: The N- and C-terminal domains of DDX3X are required for its activity. Kinetic ATPase reactions were performed in the presence of 100 nM $DDX3_{1-662}$ or of $DDX3_{132-607}$ and 100 nM (A) or 1 µM (B-D) substrate RNA. (E) 16 bp RNA/RNA duplexes (D3, 5 nM) were incubated in the presence of increasing concentrations of $DDX3_{132-607}$ -DQAD (0 - 4000 nM) and 2 mM ATP. The quantitation of bound duplexes allows for the determination of the following binding parameters: Bmax = 1.00 ± 0.04, h = 1.4 ± 0.1 and Kd = 138 ± 12 nM. (F) Representation of the results obtained in Table 3. 5 nM of RNA were incubated in the presence of 0-4000 nM recombinant protein and 2 mM ATP as described in the Supplemental Material and Methods section. The values represent the average and standard deviation of at least three independent experiments. p indicates the P value issued from an unpaired t-test whenever appropriate.

Evaluation of the requirement for DDX3X terminal extensions

A previous study performed on a truncated version of DDX3X (aa 132-607) that removes the N- and C-terminal intrinsically disordered domains of DDX3 indicated that this mutant only retains partial helicase activity [37]. To perceive the influence of these regions located outside the helicase core on the biochemical properties of DDX3, we produced and purified a MBP-DDX3$_{132\text{-}607}$ version of the protein as well as the corresponding DQAD mutant (Figure 1A). In stark contrast with the full length DDX3, the ATPase activity of the truncated DDX3$_{132\text{-}607}$ is not significantly stimulated in the presence of 100 nM RNA HIV-1 as compared to the full-length version of the protein. RNA HIV-1 concentration has to be increased up to 1 µM to observe a significant ATPase activity with DDX3$_{132\text{-}607}$. In the presence of short synthetic substrates (D3 duplex, RNA/RNA duplex with a 3'-tail) DDX3$_{132\text{-}607}$ shows a weak activity comparable to what is observed with the full length protein (Figure 7C-D). Consistent with previous results, DDX3 truncated version does not show any significant helicase activity on any of these model substrates (35, not shown). Taken together, these results show that the helicase core of DDX3 is not sufficient to sustain robust ATPase and helicase activities of DDX3. Finally, in order to evaluate the binding capacities of the truncated version of the protein, we performed gel shift assays. Similarly to what is observed with the full-length protein, we were not able to detect a significant amount of high molecular weight complex with the WT version of DDX3$_{132\text{-}607.}$ In stark contrast, DDX3$_{132\text{-}607}$-DQAD in presence of ATP forms stable complexes with the model substrates, although with a significantly reduced affinity (K_d = 138 nM). With this protein, three to four high molecular weight complexes are observed, suggesting a stoichiometry of 3-4 per substrate (Figure

7E). The reduced affinity holds true for HIV-1 derived RNAs (Figure 7F, Table 3) with a consistent 10-fold increase in the observed Kds. These results indicate that the flanking N and C-terminal regions of DDX3 are also required for DDX3 interaction with its RNA substrate.

DISCUSSION

hDDX3X is multi-functional both in the cell and during HIV-1 replication, as a consequence the understanding of its molecular role(s) necessitates the development of an *in vitro* system to independently study separate functions. In an attempt to elucidate further the role of DDX3 in HIV-1 gRNA translation initiation and to initiate the molecular dissection of this mechanism, we first overexpressed hDDX3X in *E. coli* and performed the characterization of its biochemical properties. It is important to note that in the absence of the MBP-module, isolated full-length DDX3 is poorly soluble and hardly amenable to extensive biochemical assays. In spite of these difficulties, our purification scheme allows us to produce mg amounts of an active protein. For instance, we performed ATPase assays [20,30] and confirm that DDX3 is active and that this activity relies on the integrity of its DEAD domain. However, we and others [31] could not reproduce a basal ATPase activity of DDX3 in absence of nucleic acids, nor its stimulation by dsDNA; which are controversial properties for DEAD-box proteins. Short nucleic acids substrates only induce a weak stimulation of the ATPase activity and we confirm the impossibility to reach saturation even in presence of up to 10 µM of such substrates. In contrast, the stimulation of the ATPase activity saturates in presence of 200 – 400 nM of the 1636 nucleotide long fragment of HIV-1 gRNA. We were then able to determine the kinetic parameters of DDX3 ATPase activity, and obtained a

significantly increased Kcat/K_M value of 1.12 µM. min^{-1} as compared to the values reported by Garbelli et. al. or by Sharma et al. (0.07 and 0.08 µM.min^{-1} respectively). The K_{cat}/K_M ratio is a measure of the catalytic efficiency; nonetheless, it can mask differences in individual parameters, which could result from the quality/stability of the produced protein, or probably more relevant from the nature of the nucleic acid used to stimulate DDX3 ATPase activity.

To establish that our recombinant protein is an active helicase, we used synthetic duplexes and confirm the requirement for a single-stranded region for the enzyme to be active. This is in agreement with the current model of "local strand separation" and enzyme loading on single-stranded nucleic acids [27]. Nonetheless, DDX3 displayed a preference for substrates presenting a 3' tail as previously reported [31]. Increase in the helicase activity on substrates presenting a 3' overhang could result from a differential substrate binding or from any of the steps involved in ATP hydrolysis, in helicase activity or in the coupling of both activities. Although we could not reliably measure the ATPase activity triggered by the synthetic helicase substrates, in one case we determined that increased helicase activity is associated to a decreased binding to nucleic acids. This may appear counter-intuitive but one has to remember that the Kd reflects the interaction of DDX3 with its substrate which is only the first step of the helicase activity that is the result of many individual steps as modeled for Ded1p for instance [27]. In such a scheme, many scenarios could explain our results, for example, the increased stability of the enzyme/substrate could render the duplex harder to destabilize. Alternatively, the reaction product resulting from unwinding of a 3’ overhang substrate may have more affinity for the enzyme than the product of a 5’ overhang substrate.

In contrast to what is observed with Ded1p we could not detect any RNA/DDX3 complexes when using non-hydrolysable ATP analogs. We had to resort to the DQAD mutant impaired for ATP hydrolysis to observe RNA-protein complexes in presence of ATP. Complex formation with the DDX3 DQAD is strongly reduced in presence of AMP-PNP and ATP-γS indicating that a stable nucleotide-DDX3-RNA ternary complex strictly requires ATP. Modification of ATP by a non-hydrolysable analog is not trivial as shown by structural analysis of myosin complexed with AMP-PNP which revealed that the presence of the bridging nitrogen in AMP-PNP change the hydrogen bonding pattern and could explain the lowered affinity [38].

When testing synthetic substrates of different sequences, we could not observe any difference in the stimulation of ATP hydrolysis. Since we tested only a very small set of RNAs, it is possible that we did not find the correct sequence to specifically activate DDX3. One can also envisage that DDX3, as most DEAD-box helicases, displays a low sequence specificity towards its RNA substrate and has to be promiscuous to fulfill its multiple cellular functions. To our knowledge, the bacterial DbpA is the only DEAD-box helicase which activity strictly relies on the anchoring on a specific sequence which in the present case consists in the hairpin 92 of the 23S RNA [39]. Instead, our results with a biological substrate rather suggest the presence of determinants in the A_{336}-U_{851} region of Gag ORF. Counter-intuitively, adding extra sequences to this domain seem to decrease DDX3 stimulation, which could reflect the formation of alternative structures, or result from a non-productive titration of DDX3. Deleting additional sequences to this domain allows us to define a minimal A_{336}-C_{540} fragment that retains significant capacity to stimulate DDX3 ATPase activity provided the presence of the 336-406 and 477-540

regions. Such a bipartite signal could be reminiscent of a structural element involving sequences in each of the above-mentioned regions. Importantly and to our knowledge, such structural element would match the P2 domain of HIV-1 RNA that is also involved in a recently described three-way junction [33]. Taken together, our results strongly suggest that structural determinants rather than sequences are involved in DDX3X interaction with HIV-1 RNA.

In this work, we have identified HIV-1 gRNA as a preferred substrate which strongly stimulates DDX3 ATPase activity as compared to the model substrates used so far. We were able to narrow down to the A_{336}-C_{851} region of the molecule that contains the Gag-IRES. This region could conceal a specific motif yet to be characterized or else, the specificity could be defined by the presence of alternating structures. The rationale for this peculiar specificity could reside in DDX3 function in the cell such as selectively facilitating the translation of structured mRNAs [16] or selectively interacting with G-quadruplex containing RNAs [40]. In most DEAD-box proteins, the core is flanked by additional N and C-terminal extensions, which contribute to the functional diversity of this protein family. Many of these extensions direct individual DEAD-box proteins to their functional targets by interaction with protein or RNA components of the target and some extensions modulate the activity of the helicase core [41,42]. Our results suggest that HIV-1 RNA is one of these targets, able to optimally stimulate DDX3 enzymatic activity provided the presence of the N- and C-terminal extensions. DDX3 has been described in complex with various translation initiation factors, but does stimulate the translation of only a subset of cellular or viral mRNAs. This helicase may therefore combine intrinsic

specificity for some structural motif(s) and interaction with other proteins to act on different but yet specific substrates.

SUPPLEMENTARY MATERIAL

Supplemental_Fig_S1: MBP-DDX3 purification

Supplemental_Fig_S2: MBP-DQAD purification

Supplemental_Fig_S3: Kinetic analysis of MBP-DDX3 ATPase activity

Supplemental_Fig_S4: MBP-DDX3 helicase activity

Supplemental_Table_S1: Sequence of the nucleic acids used in this study

Supplemental_Methods

ACKNOWLEDGEMENTS

The authors would like to thank Josette Banroques and Kyle Tanner for the kind gift of purified Ded1p and fruitful discussions.

FUNDING

Research in B.S. laboratory was financed by the CNRS, the University Paris Descartes and a grant from *la Fondation pour la Recherche Médicale* (FRM #DBI20141231337). G. B. and M.A are recipient of a PhD fellowship from the French ministry for research and education (Allocation MRE).

REFERENCES

[1] B. Gilman, P. Tijerina, R. Russell, Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs, Biochem. Soc. Trans. 45 (2017) 1313–1321. doi:10.1042/BST20170095.

[2] A.E. Gorbalenya, E.V. Koonin, Helicases: amino acid sequence comparisons and structure-function relationships, Curr. Opin. Struct. Biol. 3 (1993) 419–429. doi:10.1016/S0959-440X(05)80116-2.

[3] P. Linder, P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier, P.P. Slonimski, Birth of the D-E-A-D box, Nature. 337 (1989) 121–122. doi:10.1038/337121a0.
[4] M.E. Fairman-Williams, U.-P. Guenther, E. Jankowsky, SF1 and SF2 helicases: family matters, Curr. Opin. Struct. Biol. 20 (2010) 313–324. doi:10.1016/j.sbi.2010.03.011.
[5] S. Rocak, P. Linder, DEAD-box proteins: the driving forces behind RNA metabolism, Nat. Rev. Mol. Cell Biol. 5 (2004) 232–241. doi:10.1038/nrm1335.
[6] P. Linder, Dead-box proteins: a family affair—active and passive players in RNP-remodeling, Nucleic Acids Res. 34 (2006) 4168–4180. doi:10.1093/nar/gkl468.
[7] P. Linder, E. Jankowsky, From unwinding to clamping — the DEAD box RNA helicase family, Nat. Rev. Mol. Cell Biol. 12 (2011) 505–516. doi:10.1038/nrm3154.
[8] D. Sharma, E. Jankowsky, The Ded1/DDX3 subfamily of DEAD-box RNA helicases, Crit. Rev. Biochem. Mol. Biol. 49 (2014) 343–360. doi:10.3109/10409238.2014.931339.
[9] Y. Ariumi, Multiple functions of DDX3 RNA helicase in gene regulation, tumorigenesis, and viral infection, Front. Genet. 5 (2014). doi:10.3389/fgene.2014.00423.
[10] R. Soto-Rifo, T. Ohlmann, The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism, Wiley Interdiscip. Rev. RNA. 4 (2013) 369–385. doi:10.1002/wrna.1165.
[11] L. Zhao, Y. Mao, J. Zhou, Y. Zhao, Y. Cao, X. Chen, Multifunctional DDX3: dual roles in various cancer development and its related signaling pathways, Am. J. Cancer Res. 6 (2016) 387–402.
[12] M. Schröder, Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation?, Biochem. Soc. Trans. 39 (2011) 679–683. doi:10.1042/BST0390679.
[13] F. Valiente-Echeverría, M.A. Hermoso, R. Soto-Rifo, RNA helicase DDX3: at the crossroad of viral replication and antiviral immunity, Rev. Med. Virol. 25 (2015) 286–299. doi:10.1002/rmv.1845.
[14] V.S.R.K. Yedavalli, C. Neuveut, Y. Chi, L. Kleiman, K.-T. Jeang, Requirement of DDX3 DEAD Box RNA Helicase for HIV-1 Rev-RRE Export Function, Cell. 119 (2004) 381–392. doi:10.1016/j.cell.2004.09.029.
[15] A. Fröhlich, B. Rojas-Araya, C. Pereira-Montecinos, A. Dellarossa, D. Toro-Ascuy, Y. Prades-Pérez, F. García-de-Gracia, A. Garcés-Alday, P.S. Rubilar, F. Valiente-Echeverría, T. Ohlmann, R. Soto-Rifo, DEAD-box RNA helicase DDX3 connects CRM1-dependent nuclear export and translation of the HIV-1 unspliced mRNA through its N-terminal domain, Biochim. Biophys. Acta BBA - Gene Regul. Mech. 1859 (2016) 719–730. doi:10.1016/j.bbagrm.2016.03.009.
[16] R. Soto-Rifo, P.S. Rubilar, T. Limousin, S. de Breyne, D. Décimo, T. Ohlmann, DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs: Translation initiation mediated by DDX3, EMBO J. 31 (2012) 3745–3756. doi:10.1038/emboj.2012.220.
[17] R. Soto-Rifo, P.S. Rubilar, T. Ohlmann, The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA, Nucleic Acids Res. 41 (2013) 6286–6299. doi:10.1093/nar/gkt306.

[18] N.T. Parkin, E.A. Cohen, A. Darveau, C. Rosen, W. Haseltine, N. Sonenberg, Mutational analysis of the 5' non-coding region of human immunodeficiency virus type 1: effects of secondary structure on translation., EMBO J. 7 (1988) 2831–2837.
[19] N. Chamond, N. Locker, B. Sargueil, The different pathways of HIV genomic RNA translation, Biochem. Soc. Trans. 38 (2010) 1548–1552. doi:10.1042/BST0381548.
[20] I. Iost, M. Dreyfus, P. Linder, Ded1p, a DEAD-box Protein Required for Translation Initiation in *Saccharomyces cerevisiae* , Is an RNA Helicase, J. Biol. Chem. 274 (1999) 17677–17683. doi:10.1074/jbc.274.25.17677.
[21] H.A. Bowers, P.A. Maroney, M.E. Fairman, B. Kastner, R. Lührmann, T.W. Nilsen, E. Jankowsky, Discriminatory RNP remodeling by the DEAD-box protein DED1, RNA. 12 (2006) 903–912. doi:10.1261/rna.2323406.
[22] M.E. Fairman, P.A. Maroney, W. Wang, H.A. Bowers, P. Gollnick, T.W. Nilsen, E. Jankowsky, Protein Displacement by DExH/D "RNA Helicases" Without Duplex Unwinding, Science. 304 (2004) 730–734. doi:10.1126/science.1095596.
[23] F. Liu, A.A. Putnam, E. Jankowsky, DEAD-Box Helicases Form Nucleotide-Dependent, Long-Lived Complexes with RNA, Biochemistry. 53 (2014) 423–433. doi:10.1021/bi401540q.
[24] A.A. Putnam, E. Jankowsky, AMP Sensing by DEAD-Box RNA Helicases, J. Mol. Biol. 425 (2013) 3839–3845. doi:10.1016/j.jmb.2013.05.006.
[25] Q. Yang, E. Jankowsky, ATP- and ADP-Dependent Modulation of RNA Unwinding and Strand Annealing Activities by the DEAD-Box Protein DED1 †, Biochemistry. 44 (2005) 13591–13601. doi:10.1021/bi0508946.
[26] Q. Yang, M.D. Campo, A.M. Lambowitz, E. Jankowsky, DEAD-Box Proteins Unwind Duplexes by Local Strand Separation, Mol. Cell. 28 (2007) 253–263. doi:10.1016/j.molcel.2007.08.016.
[27] A.A. Putnam, Z. Gao, F. Liu, H. Jia, Q. Yang, E. Jankowsky, Division of Labor in an Oligomer of the DEAD-Box RNA Helicase Ded1p, Mol. Cell. 59 (2015) 541–552. doi:10.1016/j.molcel.2015.06.030.
[28] S.H. Park, S.G. Lee, Y. Kim, K. Song, Assignment of a human putative RNA helicase gene, DDX3, to human X chromosome bands p11.3-->p11.23, Cytogenet. Cell Genet. 81 (1998) 178–179. doi:10.1159/000015022.
[29] R. Franca, A. Belfiore, S. Spadari, G. Maga, Human DEAD-box ATPase DDX3 shows a relaxed nucleoside substrate specificity, Proteins Struct. Funct. Bioinforma. 67 (2007) 1128–1137. doi:10.1002/prot.21433.
[30] A. Garbelli, S. Beermann, G. Di Cicco, U. Dietrich, G. Maga, A motif unique to the human DEAD-box protein DDX3 is important for nucleic acid binding, ATP hydrolysis, RNA/DNA unwinding and HIV-1 replication, PloS One. 6 (2011) e19810. doi:10.1371/journal.pone.0019810.
[31] D. Sharma, A.A. Putnam, E. Jankowsky, Biochemical Differences and Similarities between the DEAD-Box Helicase Orthologs DDX3X and Ded1p, J. Mol. Biol. 429 (2017) 3730–3742. doi:10.1016/j.jmb.2017.10.008.
[32] J. Angulo, N. Ulryck, J. Deforges, N. Chamond, M. Lopez-Lastra, B. Masquida, B. Sargueil, LOOP IIId of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex, Nucleic Acids Res. 44 (2016) 1309–1325. doi:10.1093/nar/gkv1325.
[33] J. Deforges, S. de Breyne, M. Ameur, N. Ulryck, N. Chamond, A. Saaidi, Y. Ponty, T. Ohlmann, B. Sargueil, Two ribosome recruitment sites direct multiple translation

events within HIV1 Gag open reading frame, Nucleic Acids Res. 45 (2017) 7382–7400. doi:10.1093/nar/gkx303.
[34] A. Pause, N. Sonenberg, Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A., EMBO J. 11 (1992) 2643–2654.
[35] C.B. Buck, X. Shen, M.A. Egan, T.C. Pierson, C.M. Walker, R.F. Siliciano, The Human Immunodeficiency Virus Type 1gag Gene Encodes an Internal Ribosome Entry Site, J. Virol. 75 (2001) 181–191. doi:10.1128/JVI.75.1.181-191.2001.
[36] N. Locker, N. Chamond, B. Sargueil, A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3, Nucleic Acids Res. 39 (2011) 2367–2377. doi:10.1093/nar/gkq1118.
[37] S.N. Floor, K.J. Condon, D. Sharma, E. Jankowsky, J.A. Doudna, Autoinhibitory Interdomain Interactions and Subfamily-Specific Extensions Redefine the Catalytic Core of the Human DEAD-box Protein DDX3, J. Biol. Chem. (2015) jbc.M115.700625. doi:10.1074/jbc.M115.700625.
[38] A.M. Gulick, C.B. Bauer, J.B. Thoden, I. Rayment, X-ray Structures of the MgADP, MgATPγS, and MgAMPPNP Complexes of the *Dictyostelium discoideum* Myosin Motor Domain [†,‡], Biochemistry. 36 (1997) 11619–11628. doi:10.1021/bi9712596.
[39] C.M. Diges, O.C. Uhlenbeck, Escherichia coli DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA, EMBO J. 20 (2001) 5503–5512. doi:10.1093/emboj/20.19.5503.
[40] B. Herdy, C. Mayer, D. Varshney, G. Marsico, P. Murat, C. Taylor, C. D'Santos, D. Tannahill, S. Balasubramanian, Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts, Nucleic Acids Res. (2018). doi:10.1093/nar/gky861.
[41] O. Cordin, J. Banroques, N.K. Tanner, P. Linder, The DEAD-box protein family of RNA helicases, Gene. 367 (2006) 17–37. doi:10.1016/j.gene.2005.10.019.
[42] M. Hilbert, A.R. Karow, D. Klostermeier, The mechanism of ATP-dependent RNA unwinding by DEAD box proteins, Biol. Chem. 390 (2009). doi:10.1515/BC.2009.135.

TABLE 1: DDX3-DQAD binding affinities to HIV-1 gRNA fragments. Filter binding assays were performed in the presence of 5 nM RNAs and increasing concentrations of DDX3-DQAD. The results are the mean and standard deviation of two to three independent experiments. The normalization per nucleotide is shown to assess the size effect.

RNA	Nucleotides	Kd (nM) Mean ± s.e.m.	Length (nt)	Kd per nt ($pM.nt^{-1}$)
FL	G_1-U_{1636}	30.5 ± 4.8	1636	18.6
UTR	G_1-A_{416}	89.4 ± 14.7	416	215.0
Gag ORF	A_{336}-U_{1636}	9.6 ± 3.6	1301	7.4
	G_{690}-U_{1636}	37.1	947	39.2
Gag IRES	A_{336}-U_{851}	32.2 ± 5.8	516	62.4
	A_{336}-A_{715}	37.1 ± 5.0	380	97.6
	A_{336}-G_{656}	60.7 ± 8.5	321	189.1
	A_{336}-C_{540}	54.3 ± 17.6	205	264.9
	G_{406}-C_{540}	> 500	135	-
	A_{336}-C_{477}	59.2 ± 12.7	152	389.5
Δ Site1 (Δ1)	A_{336}-U_{881}Δ363-485	39.0 ± 11.4	1178	33.1
Δ Site2 (Δ2)	A_{336}-U_{881}Δ632-756	43.2 ± 8.2	1195	36.2
CTRL 1		63.0 ± 29.6	1031	61.1
CTRL 2		31.8 ± 5.2	505	63.0
CTRL 3		60.8 ±31.6	243	250.2
CTRL 4		106.8 ±34.8	183	583.6
Duplex D3		> 500	57	-
RNA L		> 500	41	-

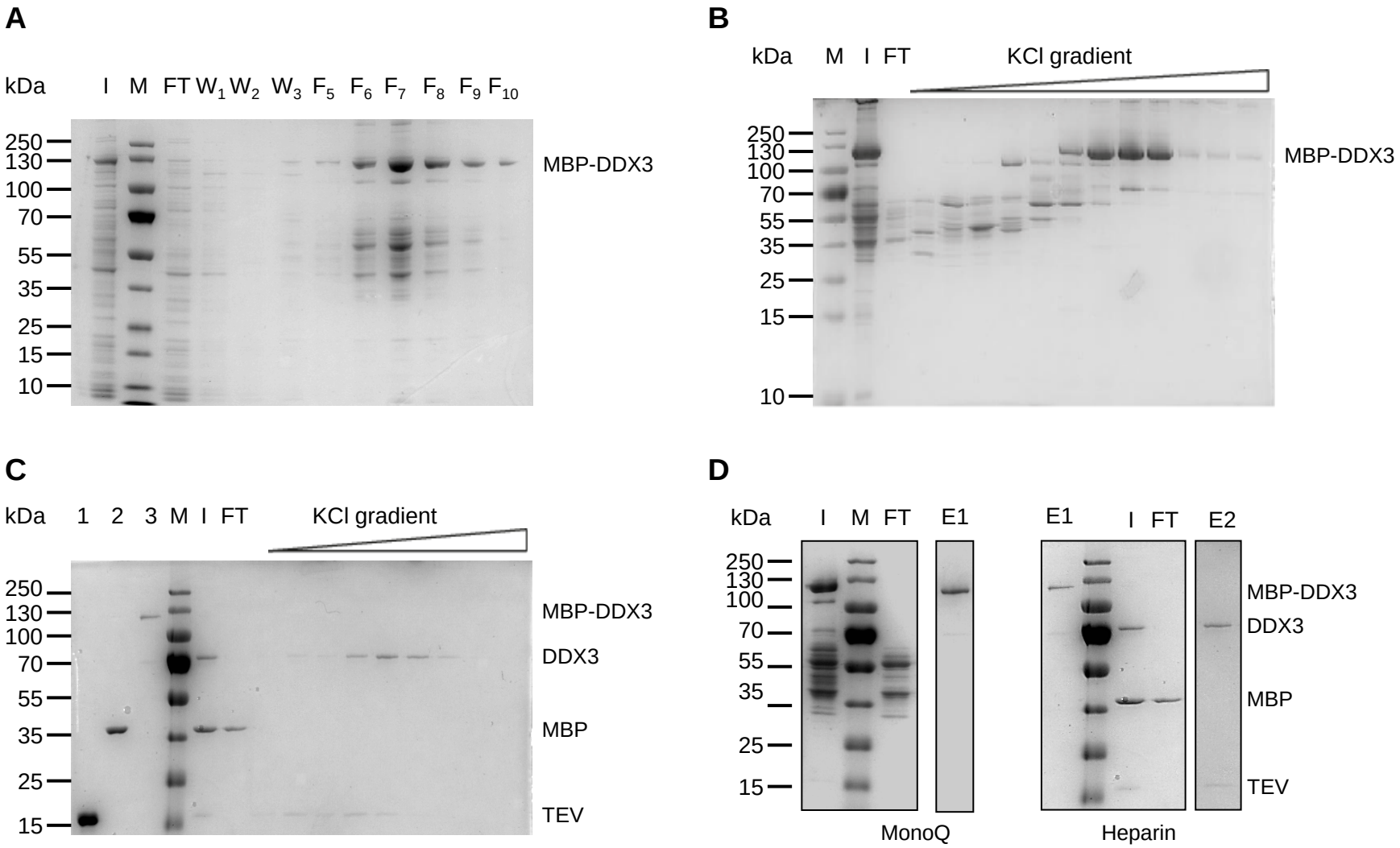
TABLE 2: DDX3-WT ATPase activity after stimulation by HIV-1 derived RNAs. Kinetic ATPase assays were performed with 100 nM DDX3, 100 nM RNA and 10 µM ATP. The results are the mean and standard deviation of at least three independent experiments. The normalization per nucleotide is shown to assess the size effect.

RNA	Nucleotides	Vi ($nM.min^{-1}$)	Length (nt)	Vi per nt ($pM.min^{-1}.nt^{-1}$)
FL	G_1-U_{1636}	85 ± 7	1636	52.0
UTR	G_1-A_{416}	30 ± 4	416	72.1
Gag ORF	A_{336}-U_{1636}	105 ± 6	1301	80.7
	G_{690}-U_{1636}	48.6	947	51.3
Gag IRES	A_{336}-U_{851}	152 ± 6	516	294.6
	A_{336}-A_{715}	132 ± 3	380	347.4
	A_{336}-G_{656}	86 ± 9	321	267.9
	A_{336}-C_{540}	73 ± 8	205	356.1
	G_{406}-C_{540}	5 ± 1	135	37.0
	A_{336}-C_{477}	14 ± 4	152	92.1
Δ Site1 (Δ1)	A_{336}-U_{881}Δ363-485	109	1178	92.5
Δ Site2 (Δ2)	A_{336}-U_{881}Δ632-756	127	1195	106.3
Δ1 Δ2	A_{336}-U_{881}Δ363-485Δ632-756	122	1072	113.8
CTRL 1		27.5	1031	26.7
CTRL 2		48.4	505	95.8
CTRL 3		4	243	16.5
CTRL 4		4	183	21.9
Duplex D3		5 ± 1	57	87.7

TABLE 3: $DDX3_{132-607}$ DQAD binding affinities to HIV-1 gRNA fragments. Filter binding assays were performed in the presence of 5 nM RNAs and increasing concentrations of DDX3-DQAD. The results are the mean and standard deviation of two to three independent experiments. The normalization per nucleotide is shown to assess the size effect.

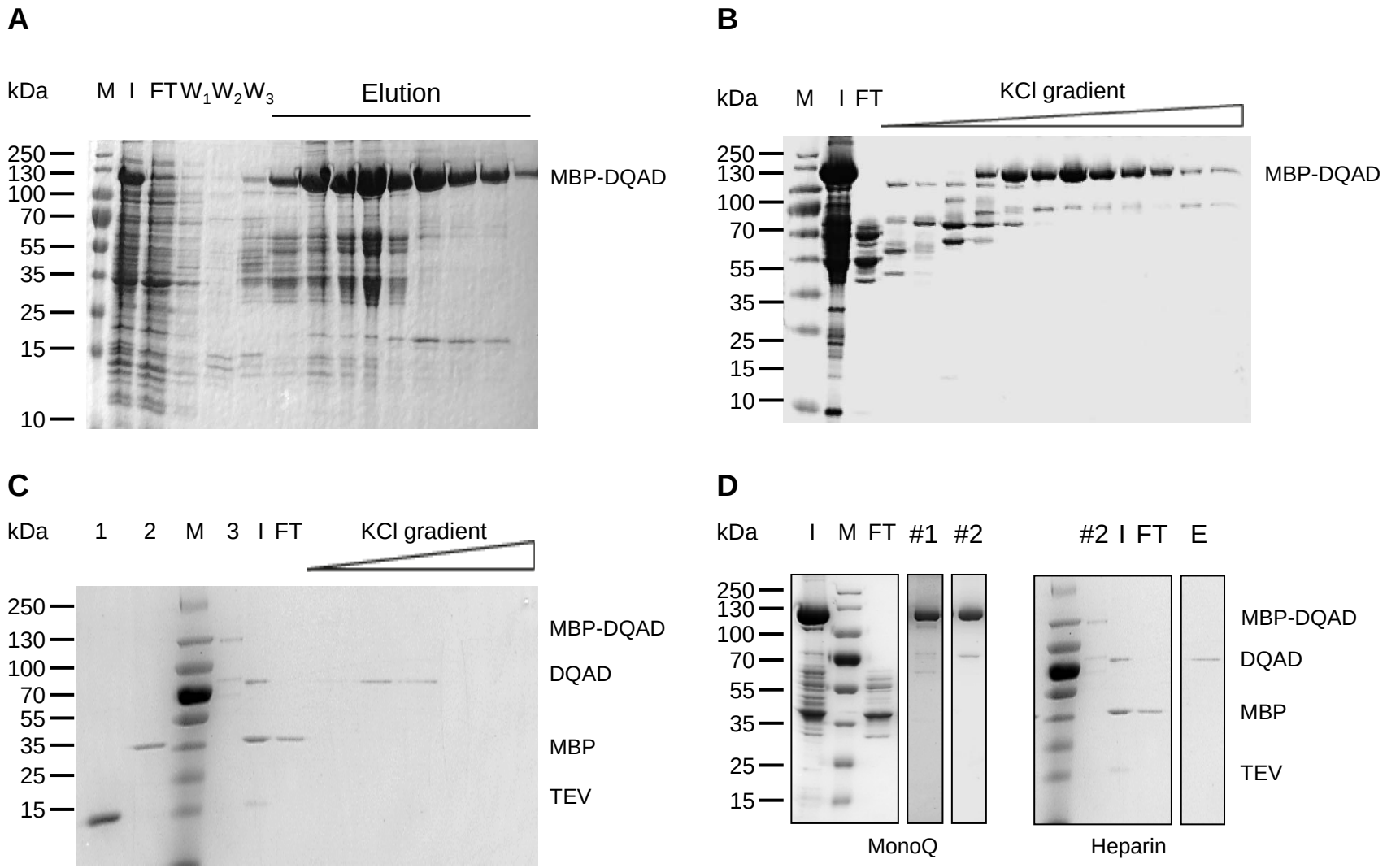
RNA	Nucleotides	Kd (nM) Mean ± s.e.m.	Length (nt)	Kd per nt (pM/nt)
FL	G_1-U_{1636}	183 ± 24	1636	112
UTR	G_1-A_{416}	473 ± 90	416	1137
Gag ORF	A_{336}-U_{1636}	462 ± 113	1301	355
Gag IRES	A_{336}-U_{851}	423 ± 116	516	820
	A_{336}-A_{715}	270 ± 130	380	711
	A_{336}-C_{540}	633 ± 157	205	3088
	G_{406}-C_{540}	870 ± 187	135	6444
	A_{336}-C_{477}	510 ± 133	152	3355

Supplementary Figure S1



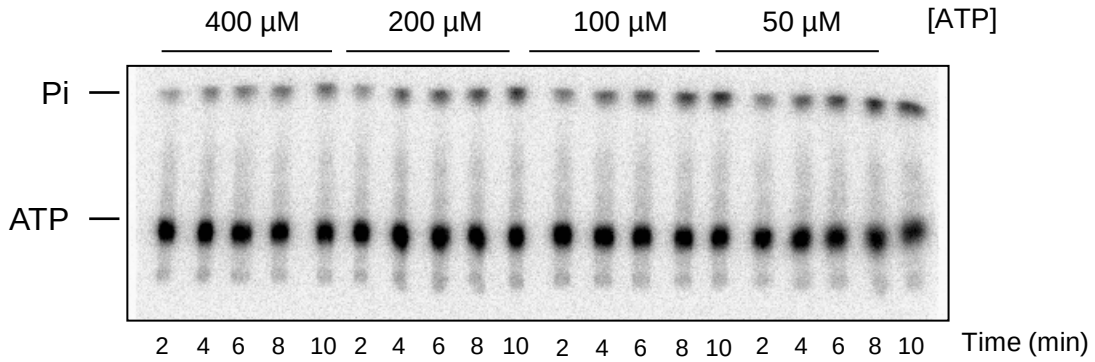
Supplementary Figure S1: **Purification of MBP- DDX3**. SDS-Page analysis of (**A**) MBP-DDX3 purification with a Ni-NTA column. W_1 to W_3, washes performed with 10 mM (W_1 and W_2) to 60 Imidazole (W_3) and 1 M KCl (W_1) ; F_5 to F_{10}, eluted fractions. (**B**) F_5 to F_{10} from (A) separation with a MonoQ over a linear KCl gradient (0 - 1 M KCl). (**C**) The protein was TEV cleaved and purified with a heparin column over a linear KCl gradient. Loading controls include His-TEV (Lane 1), His-MBP (Lane 2) and uncleaved MBP-DDX3 (Lane 3). (**D**) Summary of MBP-DDX3 purification. I, input ; FT, flow-through. Molecular weight standards (M, kDa) are indicated in the left margin.

Supplementary Figure S2



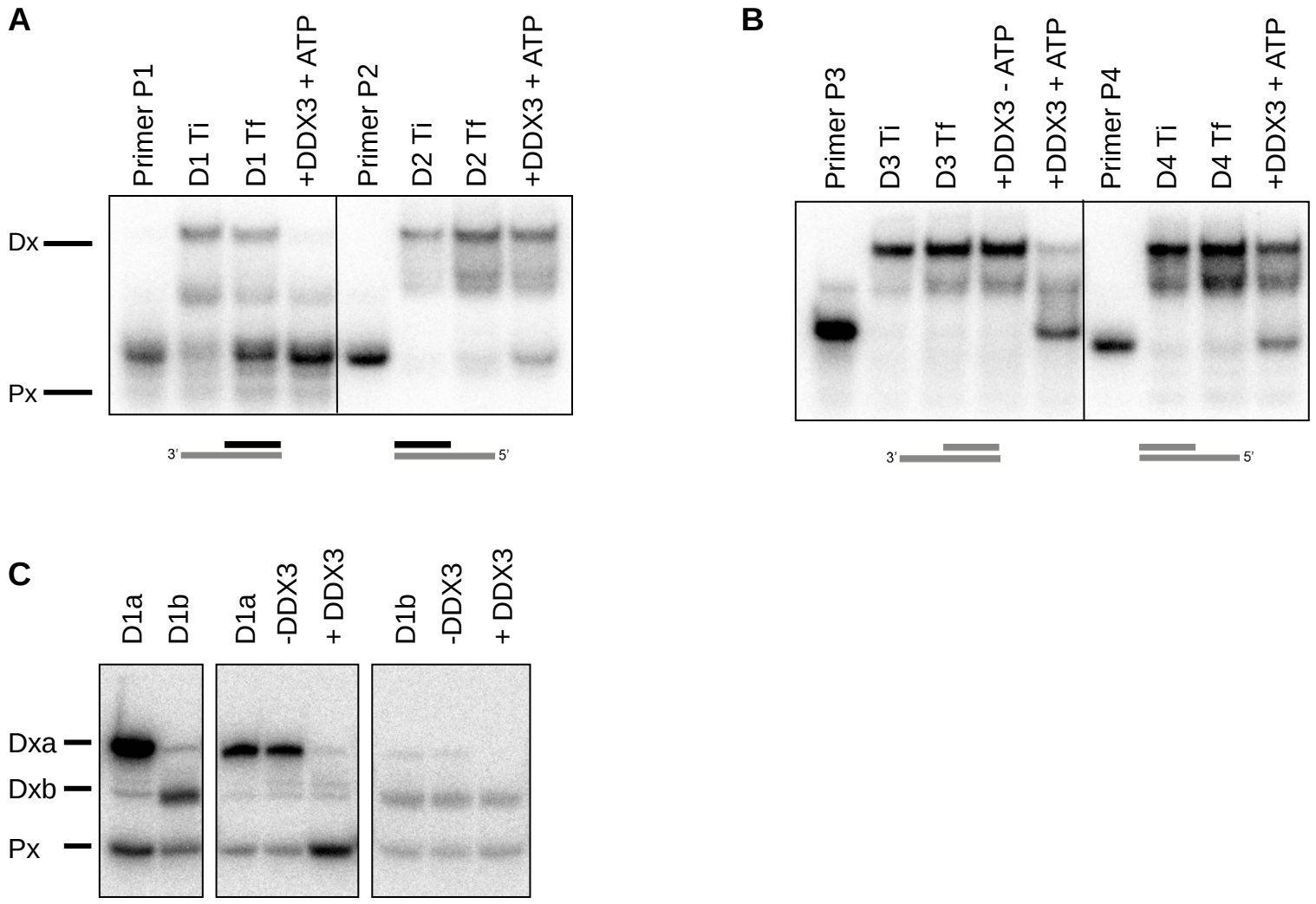
Supplementary Figure S2: **SDS-PAGE analysis of MBP-DQAD purification**. (**A**) MBP-DQAD purification with a Ni-NTA column. W_1 to W_3, washes performed with 10 mM (W_1 and W_2) to 60 Imidazole (W_3) and 1 M KCl (W_1) ; Elution, F_3 to F_{10} and F_{20} fractions eluted with 0.5 M KCl. (**B**) F_4 to F_{10} from (A) separation with a MonoQ over a linear KCl gradient (0 - 1 M KCl). (**C**) The protein was TEV-cleaved and purified with a heparin column over a linear KCl gradient. Loading controls include His-TEV (Lane 1), His-MBP (Lane 2) and uncleaved MBP-DQAD (Lane 3). (**D**) Summary of MBP-DQAD purification. I, input ; FT, flow-through. Molecular weight standards (M, kDa) are indicated in the left margin.

Supplementary Figure S3



Supplementary Figure S3: **Kinetic analysis of MBP-DDX3 ATPase activity**. MBP-DDX3 (100 nM) was incubated in the presence of 100 nM HIV-1 RNA (1636 nt) and various concentrations of ATP containing traces of α^{32}P-ATP. Aliquotes were removed every 2 minutes and deposited onto a PEI cellulose plate. After migration, the proportion of Pi released (as of total ATP) was quantified for the different time points and allows for Vi determination at different ATP concentrations.

Supplementary Figure S4



Supplementary Figure S4: **Helicase activity of DDX3**. 5 nM of DNA/RNA (**A**) or RNA/RNA duplexes (**B**) were incubated in the presence of 140 nM of purified enzyme for 5 minutes. Relative primer release was measured as indicated in the methods section. Ti, initial duplex, Tf, duplex after 5 min at 37°C. Duplexes (as described in Figure 4A) are schematically represented underneath, DNA is indicated in black and RNA in grey. (**C**) Hybridization of a model substrate (D1) give rise to 2 major species (D1a and D1b) which were gel purified for helicase assays. Only D1a is a substrate for DDX3 helicase activity.

SUPLEMENTARY MATERIAL AND METHODS

In vitro transcription

RNAs were directly transcribed using the T7 RNA polymerase from polymerase chain reaction (PCR) products containing the T7 promoter sequence (5'-TAATACGACTCACTATA-3') and purified as previously described (1). Radiolabeled RNAs were transcribed as above, in the presence of α^{32}-UTP (3000 mCi/mmol) and purified by size-exclusion chromatography.

Mobility shift assay

Various concentrations of purified protein were mixed with 2 mM ATP, AMP-PNP or ATPγS and 7 nM radiolabeled RNA:RNA or RNA:DNA substrate and incubated in unwinding buffer for 5 minutes at 37°C. Samples were then chilled on ice, stopped with a solution containing 50 mM EDTA, 0.1 % Xylene Cyanol, 0.1 % Bromophenol Blue and 20 % glycerol. Samples were loaded on 4% native polyacrylamide gels, which were run in cold for 1 hour at 100 V, dried, exposed and analyzed using a BAS-5000 scanner and the software MultiGauge.

Filter binding assay

Radio-labeled RNA (2.5-10 nM) was denatured by heating to 80°C for 2 min and then cooled to room temperature in FB buffer (20-mM Tris-Cl pH 8, 100-mM KOAc, 100-mM KCl, 2.5 mM $MgCl_2$, 2-mM DTT). Serial dilutions of recombinant protein were prepared extemporaneously, added to a 10 µl reaction and incubated at 37◦C for 20 min. The reactions were then used for filter binding assays. Filter binding was accomplished

essentially as previously described (2) using two filters: from top to bottom, a nitrocellulose filter and a charged nylon filter. The filters were presoaked in FB buffer, assembled in the dot blot apparatus and the reactions were applied and directly vacuum filtered. The filters were then rinsed with FB buffer, removed and radioactivity was quantified using a storm phosphorImager (GE Healthcare). To determine the apparent *K*d, the data were fit to the Langmuir isotherm described by the equation:

$\theta = P/[P + K\mathrm{d}]$, where θ is the fraction of RNA bound and P is the protein concentration.

1. Locker, N., Chamond, N. and Sargueil, B. (2011) A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. *Nucleic Acids Res*, **39**, 2367-2377.
2. Kieft, J.S., Grech, A., Adams, P. and Doudna, J.A. (2001) Mechanisms of internal ribosome entry in translation initiation. *Cold Spring Harb Symp Quant Biol*, **66**, 277-283.

All sequences are 5' to 3'

Substrates for ATPase assays (Figure 2)

RNA26	GGGUUUUUUAAUUUUUUAAUUUUUUC
RNA 57 (TAR)	GGUCUCUCUGGUUAGACCAGAUCUGAGCCUGGGAGCUCUCUGGCUAACUAGGGAACC
RNA64	GAACAACAACAACAACAACAGAAAAAAUUAAAAAAUUAAAAAACUCGGAGGGGCCGGUGGGGCC
Partial Duplex	RNA64/RNA26
Heteroduplex	RNA26/DNA24
DNA24	GAAAAAATTAAAAAATTAAAAAAC
DNA64	GAACAACAACAACAACAACAGAAAAAATTAAAAAATTAAAAAACTCGGAGGGGCCGGTGGGGCC
DNA70 (TAR)	TAATACGACTATAGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACC
DNA82	TAATACGACTATAGAACAACAACAACAACAACAGAAAAAATTAAAAAATTAAAAAACTCGGAGGGGCCGGTGGGGCC

Substrates for Helicase assays

RNA41_L	GCGUCUUUACGGUGCUUAAAACAAAACAAAACAAAACAAAA
RNA41_T	AAAACAAAACAAAACAAAACAAAAUAGCACCGUAAAGACGC
DNA P1	AGCACCGTAAAGACGC
DNA P2	GCGTCTTTACGGTGCT
RNA P3	AGCACCGUAAAGACGC
RNA P4	GCGUCUUUACGGUGCU
D1	P1/RNA41_L
D2	P2/RNA41_T
D3	P3/RNA41_L
D4	P4/RNA41_T
B1	RNA P4/RNA P3
B2	DNA P2/DNA P1
B3	DNA P2/RNA P4

Primers for the different RNA fragments used in this study

Fwd Primers	
G1	TAATACGACTCACTATAGGATGGGTGCGAGAGCGTCGGTCTCTCTGGTTAGACCAGATC
A336	TAATACGACTCACTATAGGATGGGTGCGAGAGCGTCGG
G406	TAATACGACTCACTATAGGGAAAGAAACAATATAAACTAAAACATATAGTATGG
G690	TAATACGACTCACTATAGGCAGCTGACACAGGAAACAACA
Reverse Primers	
A1636	AAAAAATTAGCCTGTCTCTCAGTACAATCTTTCATTTGGTGTCCT
U851	AAACATGGGTATTACTTCTGG
A715	TGGCTGTTGTTTCCTGTGTC
G656	CTCTTCCTCTATCTTATCTAAGGCTTCC
U634	AAACATGGGTATTACTTCTGG
C540	GAAGGGATGGTTGTAGCTGTCC
C477	GAAGGGATGGTTGTAGCTGTCC
A416	AAGGCCAGGGGGAAAGAAA

CTRL sequence and primers

Fwd Primer	
T7RLuc	TAATACGACTCACTATAGTGGTGCTAGCGGATCCT
Reverse Primers	
RLucas504	CAATATCTTCTTCAATATCAGGCC
RLucas243	TGGTATAATACACCGCGC
RLucas182	AGAGGCCGCGTTACC
RLucas1030	TCACTTCTTCTTCACCCGGGAG