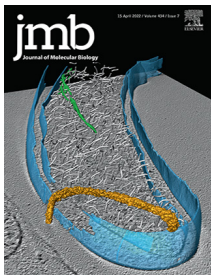
Research Article



In vitro Selection of High Affinity DNA and RNA Aptamers that Detect Hepatitis C Virus Core Protein of Genotypes 1 to 4 and Inhibit Virus Production in Cell Culture

Beatriz Torres-Vázquez [1,2], Ana María de Lucas [1], Carlos García-Crespo [3], Juan Antonio García-Martín [4], Adrián Fragoso [1†], María Fernández-Algar [1], Celia Perales [3,5,6], Esteban Domingo [3,5], Miguel Moreno [1*] and Carlos Briones [1,5*]

*1 - **Centro de Astrobiología (CAB,** CSIC-INTA), Torrejón de Ardoz, Madrid, Spain*
*2 - **Universidad de Alcalá,** Facultad de Medicina, Madrid, Spain*
*3 - **Centro de Biología Molecular Severo Ochoa (CBMSO,** CSIC-UAM), Madrid, Spain*
*4 - **Centro Nacional de Biotecnología (CNB,** CSIC), Madrid, Spain*
*5 - **Centro de Investigación Biomédica en Red de enfermedades hepáticas y digestivas (CIBERehd),** Instituto de Salud Carlos III, Madrid, Spain*
*6 - **Departamento de Microbiología Clínica,** Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain*

***Correspondence to Carlos Briones and Miguel Moreno:** Centro de Astrobiología (CAB, CSIC-INTA), Torrejón de Ardoz, Madrid, Spain. cbriones@cab.inta-csic.es @brionesci (C. Briones), morenomm@cab.inta-csic.es @migmormol (M. Moreno)*

https://doi.org/10.1016/j.jmb.2022.167501
Edited by Eric O. Freed

Abstract

Hepatitis C virus (HCV) core is a highly conserved and multifunctional protein that forms the viral capsid, making it an attractive target for HCV detection and inhibition. Aptamers are *in vitro* selected, single-stranded nucleic acids (RNA or ssDNA) with growing applicability in viral diagnostics and therapy. We have carried out DNA and RNA *in vitro* selection against six different variants of HCV core protein: two versions of the full-length protein of genotype 1, and the hydrophilic domain of genotypes 1 to 4. The aptamer populations obtained were analyzed by means of Ultra-Deep Sequencing (UDS), the most abundant sequences were identified and a number of highly represented sequence motifs were unveiled. Affinity (measured as the dissociation constant, *Kd*) of the most abundant DNA and RNA aptamers were quantified using Enzyme-Linked OligoNucleotide Assay (ELONA)-based methods. Some aptamers with nanomolar or subnanomolar *Kd* values (as low as 0.4 nM) were the common outcome of DNA and RNA selections against different HCV core variants. They were tested in sandwich and competitive biosensor assays, reaching a limit of detection for HCV core of 2 pM. Additionally, the two most prevalent and high affinity aptamers were assayed in Huh-7.5 reporter cell lines infected with HCV, where they decreased both the viral progeny titer and the extracellular viral RNA level, while increasing the amount of intracellular viral RNA. Our results suggest that these aptamers inhibit HCV capsid assembly and virion formation, thus making them good candidate molecules for the design of novel therapeutic approaches for hepatitis C.

Introduction

Hepatitis C virus (HCV) is the etiological agent of chronic hepatitis C and a major cause of fibrosis, cirrhosis and hepatocellular carcinoma.[1–5] It is estimated that around 71 million people have chronic HCV infection worldwide and up to 4 million new infections occur each year.[6–9] Approximately 30% of patients overcome the disease and the virus clears after the acute infection, while 70% remain chronically infected. Among the latter, the risk for developing cirrhosis and hepatocellular carcinoma varies between 15% and 30% within 20 years.[10] Unfortunately, a protective vaccine for HCV is not available yet.

HCV is a spherical and enveloped virus that belongs to the family *Flaviviridae* and the genus *hepacivirus.* It contains a 9.6-kb long single-stranded RNA genome with positive polarity [(+) ssRNA], composed of a 5' untranslated region (UTR), an open reading frame that encodes a single polyprotein, and a 3' UTR. Cap-independent translation starts when the ribosome of the infected cell binds to the internal ribosome entry site (IRES) element placed at the 5' UTR of the viral genome. This leads to the synthesis of a polyprotein of about 3,000 amino acids (aa), which is afterwards processed by host and viral proteases in the endoplasmic reticulum (ER) to produce three structural proteins (core, and the envelope proteins E1 and E2) as well as seven nonstructural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B).[11–12] The high mutation rate associated to the replication of HCV genome dictates the quasispecies dynamics of the virus[13] and the very large sequence variability among HCV isolates detected worldwide. HCV has been classified into seven major genotypes and several subtypes based on sequence homology, with genotypes 1 to 4 being responsible for about 95% of the reported hepatitis C infections.[14–15] In 2018, a novel HCV genotype (number 8) was identified.[16]

HCV core shows the lowest sequence variability among the 10 virus-coded proteins, with an average nucleotide and amino acid diversity of 8.72% and 4.36%, respectively.[17–18] Core is also the HCV protein which accumulates the lowest number of nucleotide mutations and amino acid substitutions during viral replication in the absence of external perturbations.[19] Such a high degree of sequence conservation suggests that HCV core is subject to relevant structural and functional restrictions. In the HCV replication cycle, core is the first protein to be translated and released by proteolytic cleavage to generate an immature, 191 aa-long polypeptide within the ER. Such a full-length core protein (known as p23 due to its molecular weight of 23 kDa) is processed in the ER by cellular proteases, giving rise to the mature form of the protein (p21), of about 177 aa in length[20] and composed of two functional domains. The first is the highly basic and hydrophilic N-terminal domain (termed D1 and spanning aa 1 to 121), which is involved in the interaction with HCV genomic RNA and in the oligomerization of the capsid around the genome. The C-terminal, hydrophobic domain (D2, aa 122 to 177), mediates core binding to ER membrane and lipid droplets (LD), which is a crucial step for virus assembly.[11] Besides forming the HCV capsid that surrounds and protects the genomic RNA, core is a multifunctional protein that interacts with cellular and viral proteins and is involved in genome transcription, lipid metabolism and cell signaling.[21–26] As a consequence of the above mentioned properties, HCV core protein is an attractive target for viral diagnosis and for the development of new antiviral drugs.

In the clinical setting, HCV infection is diagnosed by means of a two steps-protocol[27–30]: an initial serologic assay is performed to detect anti-HCV antibodies (using the third-generation enzyme immunoassay, EIA) and, if this screening test is positive, RTqPCR amplification of the viral genome is used to confirm an active case. This makes HCV diagnosis time consuming and expensive. Therefore, straightforward and cost-effective point-of-care (POC) tests with enough sensitivity are required, especially for monitoring the prevalence of the hepatitis C epidemic in developing countries as well as for massive screening campaigns. HCV structural proteins, in particular core, are suitable targets for the early diagnosis of the infection. Interestingly, HCV core can be detected in complete virions as well as in naked (full or fragmented) capsids present in the blood of infected patients.[31]

In parallel, novel drug candidates are required as alternatives to available direct-acting antivirals (DAAs), which are very efficient molecules for the treatment of hepatitis C but unaffordable options in the countries that show the highest prevalence of HCV infection. Besides, HCV viral load is not efficiently suppressed in 2% to 5% of DAA-treated patients, and antiviral drug resistance to DAAs has already been documented, with resistance-associated substitutions mainly affecting the DAA target proteins NS3, NS5A, and NS5B.[32–34] For these reasons, the development of complementary therapeutic approaches focused on alternative viral targets should be encouraged.

Aptamers are single-stranded DNA or RNA oligonucleotides that, due to their 3D structure in solution, can recognize and bind to a broad range of specific targets (including low molecular weight compounds, peptides, proteins, nucleic acids, cells or even tissues) with specificities and affinities analogous or even higher than those of antibody-antigen pairs.[35] They are selected *in vitro* using a process called SELEX (Systematic Evolution of Ligands by EXponential enrichment),[36] which consists of iterative rounds of selection-amplification of target-bound nucleic acid sequences.[37–38] Aptamers show many advantages

over antibodies: laboratory animals are not needed for their production, they have a long shelf life, display high batch consistency, can be selected against non-immunogenic or toxic targets, do not trigger any immune response, and can be chemically modified to increase their stability or functionality.[39–40] Biosensors that use aptamers as recognition elements, called aptasensors, produce a quantitative or semi-quantitative signal when specific aptamer-target interactions occur.[41] Therefore, aptamers are viable alternatives for diagnosis and therapy, with increasing applicability in virology.[42–45] Some of them have been selected against HCV targets, including a few examples of aptamers specific for HCV core protein (either the first 114 aa or the full-length protein) expressed in *E. coli*.[46–48]

In this work, we have selected DNA and RNA aptamers against 6 different variants of HCV core protein: two versions of the full-length protein of genotype 1, as well as its hydrophilic domain of genotypes 1 to 4. Five out of the 6 target proteins were expressed in insect cells. We used different SELEX protocols to find aptamers that could bind to the maximum number of variants of this key HCV protein. The obtained aptamer populations were examined by Ultra-Deep Sequencing (UDS) and bioinformatics analysis, and the affinity (measured as the dissociation constant, *Kd*) for HCV core of highly represented aptamers was assayed using Enzyme-Linked OligoNucleotide Assay (ELONA)-based methods, as previously described.[49] Subnanomolar *Kd* values (as low as 0.4 nM) showed by the best aptamers obtained, together with their excellent performance in sandwich and competitive assays, favor their use as bioaffinity probe molecules in HCV diagnosis. Alongside, the two most prevalent and high affinity aptamers were assayed in HCV infections in cell culture, where they decreased both the viral progeny titer and the extracellular viral RNA level, and concomitantly increased the presence of non-encapsidated, intracellular viral RNA. These findings anticipate the usefulness of these aptamers as efficient anti-HCV molecules.

Results

In vitro selection of DNA and RNA aptamers against HCV core protein variants

DNA and RNA aptamers were *in vitro* selected against six variants of HCV core: two versions of the full-length (191 aa long) protein of HCV genotype 1 (termed G1-191-F and G1-191) and the 122 aa-long hydrophilic domain of HCV core belonging to genotypes 1 to 4 (termed G1-122, G2-122, G3-122 or G4-122, respectively) (Supplementary Figure S1 and Supplementary Table S1). In the 12 SELEX processes performed, 10 to 14 rounds of selection-amplification were completed. Supplementary Table S2 shows a summary of the protein targets, aptamer types and selection names. A stepwise increase of the selection stringency was introduced in each round to foster the affinity of the obtained aptamers, and different counter-selection steps allowed discarding unspecific aptamers (summarized in Supplementary Table S3).

After the last round of each selection process, UDS-Illumina sequencing allowed identifying the most represented individual aptamers and the cluster of related sequences around them (see Materials and Methods). Supplementary Table S4 contains the aptamer sequences longer than 62 nt (most of them, longer than 74 nt) and shorter than 78 nt that showed a relative frequency above 100 counts per million (CPM) in at least one of the last rounds of the 12 SELEX processes performed. Figure 1 graphically depicts the relationships among the aptamers found in the last round of at least 6 of the 12 selection processes. A number of aptamers were obtained against different variants of HCV core, which were sometimes the result of both DNA and RNA SELEX processes. Such convergence in the outcome of the selection of aptamers was relevant for the aim of this work, since a subgroup of them could be useful in diagnostic and therapeutic applications for hepatitis C.

Different criteria were used to choose a subset of the highly represented aptamers for further analysis, including their presence in the last SELEX round of both DNA and RNA processes and their mean (DNA + RNA) CPM value. Four aptamers were present in a proportion over 100 CPM in at least 5 selection experiments (including at least one of DNA and one of RNA), had a mean CPM (DNA + RNA) value over 1,000 CPM, showed a very high CPM value (over 20,000) in at least one of the individual SELEX processes and/or high CPM values (over 1,000) in at least one DNA and one RNA selection (Figure 1 and Supplementary Table S4): AptD-421, AptD-1312, AptD-1932 and AptD-2251.

The aptamer AptD-1312 appeared in the last round of the 6 DNA and 6 RNA selections performed (over 100 CPM in 7 of them), showed the highest mean CPM in DNA selections (5 times higher than AptD-2251, see below), the second highest mean CPM in all selections (DNA + RNA), and the highest CPM value in a DNA process (07DS10, against the HCV core variant G1-191-F). In turn, AptD-421 was also found in the 6 DNA and 6 RNA selections performed, had the highest mean (DNA + RNA) CPM among those present in a proportion over 100 CPM in 6 selection experiments, and showed the highest CPM value in the RNA process 07RS10 (against the HCV core variant G1-191-F). AptD-2251 showed the highest mean CPM in all selections (DNA + RNA), the highest mean CPM in RNA selections (6 times higher than AptD-1312) and the highest CPM

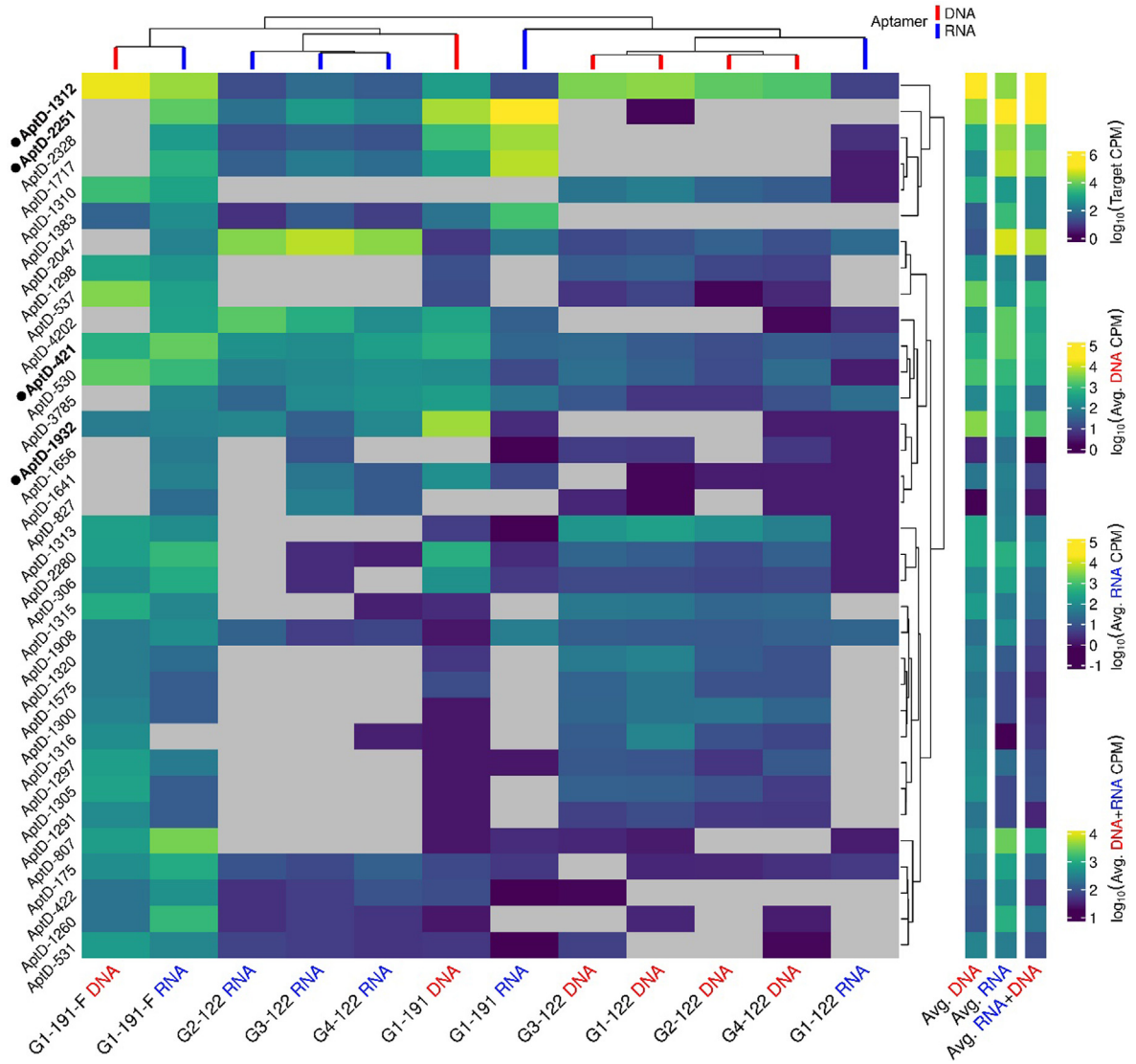


Figure 1. Heat map of the DNA and RNA aptamers found in the last round of at least 6 of the 12 SELEX processes performed. The color of the cells (using a color scale from purple to yellow) shows the $\log_{10}$ (CPM) of each aptamer in the final round (the complete list of aptamers and their CPMs is listed in Supplementary Table S4), and the average $\log_{10}$ (CPM) of DNA, RNA and DNA + RNA aptamers over all targets is shown on the right side. The grey cells correspond to the aptamers that did not appear in the final round of some of the selection processes. The order of the aptamers and targets is based on the results of a hierarchical clustering, represented by the dendrograms placed at the top and right side of the heat map. The highly represented aptamers chosen for further analysis (AptD-1312, AptD-2251, AptD-421 and AptD-1932, see text for details) are highlighted in bold and identified with a black circle.

value in an RNA process (12RS10, against the HCV core variant G1-191), although it was scarcely selected in DNA processes. Finally, AptD-1932 was obtained in 3 DNA and 6 RNA selections, showed the third highest mean CPM in DNA selections and the third highest CPM value in a DNA process (33DS10, against the HCV core variant G1-191).

Table 1 summarizes the abundance of the four highly represented aptamers in the last round of each of the 12 SELEX processes performed, and their nucleotide sequences are shown in Table 2. These aptamers, together with 22 additional ones that also met the above mentioned criteria (Supplementary Table S5) were assayed using ELONA-based methods (see below).

Secondary structure prediction and motif analysis of the highly represented aptamers

Secondary structures of the highly represented DNA and RNA aptamers were predicted using Mfold, and QGRS Mapper allowed to predict the formation of G-quadruplexes (Supplementary Figure S2). AptD-1932 was the only DNA aptamer that showed a relatively compact secondary

Table 1 Highly represented DNA and RNA aptamer sequences selected against the six variants of HCV core protein. Color code (from purple to yellow) shows the abundance of each aptamer [values in $\log_{10}$ (CPM)] in the last round of the each of the 12 SELEX processes performed (see Figure 1, Supplementary Table S4 and main text for details). The grey cells correspond to the aptamers that did not appear in the final round of some of the selection processes. The most widespread aptamer sequence was AptD-1312. Aptamers AptD-1932, AptD-2251 and AptD-421 were present in different DNA and/or RNA processes (see Figure 1), though they were mainly selected against the full-length HCV core protein (G1-191-F and G1-191).

		AptD-1312	AptD-1932	AptD-2251	AptD-421
G1-191-F	DNA	4.95	2.19		3.24
	RNA	4.33	2.33	3.85	3.9
G1-191	DNA	2.86	4.37	4.44	3.23
	RNA	1.29	0.72	5.09	1.75
G1-122	DNA	4.23		0.16	1.53
	RNA	1.1	0.49		1.4
G2-122	DNA	3.85			1.29
	RNA	1.29	2.4	1.98	2.65
G3-122	DNA	4.11			1.79
	RNA	1.84	1.63	2.78	2.5
G4-122	DNA	3.7	0.5		1.52
	RNA	1.55	2.44	2.38	2.88

structure at 37 °C under ionic conditions similar to those used during SELEX experiments, while the other three aptamers were predicted to adopt very open structures. In turn, RNA aptamers showed values of the free energy (ΔG) associated to their minimum free energy (MFE) structures much lower than those of the DNA aptamers, though it must be noticed that the only ionic conditions allowed by Mfold for RNA sequences were 1 M Na^+ and no divalent cations. G-quadruplexes, that could stabilize the tertiary structure of G-rich sequences in solution, were predicted to form in AptD-1932, AptD-2251 and the four RNA aptamers.

Cluster analysis of the aptamer populations found in the last round of the 12 SELEX processes revealed that the sequence variability was smaller in DNA aptamers than in RNA aptamers. A network of the most represented aptamers found in the last round of the DNA selection processes was constructed, in which two aptamers are connected if they share any of the five most frequent 6-mer motifs in their variable, central region plus two flanking nucleotides at each side (Figure 2(A) and Supplementary Table S6). This network analysis showed that AptD-1312 is the most connected node: it belongs to a cluster of highly interconnected aptamers (which includes AptD-1932 and AptD-2251) and also shares specific motifs with other sequences that were only selected against some of the HCV core targets. The most frequent motif found in the selected DNA aptamers against all targets is tTTGTTc (upper case letters are used for conserved nucleotides, while lower case letters show the most represented nucleotide in the variable positions of the motif), followed by tgTAAT and KCCCTc (K stands for either T or G nucleotide at the first position of the 6-mer motif). Indeed, AptD-2251, which is highly represented in selections against target G1-191, contains two copies of the tTTGTTc motif and many other GT-rich regions, but neither tgTAAT nor KCCCTc motifs. On the other hand, AptD-421 is not directly connected to AptD-1312 in the 6-mer networks, though both of them share several 5-mers (not shown).

Two different topologies were identified in the 6-mer motif networks of DNA aptamers selected against 122 or 191 aa long HCV core targets, suggestive of more than one binding mechanism: strongly connected networks with large motif similarity among all members for the hydrophilic domains of HCV genotypes 1 to 4 (Supplementary Figure S3(A)), or hybrid topologies for G1-191 (Supplementary Figure S3(B)) and G1-191-F (not shown). Interestingly, all the aforementioned motifs are predicted to be exposed in the AptD-1312 secondary structure (Supplementary Figure S2).

On the other hand, the RNA version of aptamer AptD-1312 is not among the most connected nodes (Figure 2(B)). Moreover, the weight of the edges of the network (which is proportional to the number of shared 6-mer motifs) evidences lower similarity among the RNA sequences with respect to the DNA ones. This is not surprising, due to the

Table 2 Nucleotide sequences of the four highly represented aptamers chosen for chemical synthesis and further analysis, both in DNA ("AptD-") and RNA ("AptR-") versions. The central, variable region of each sequence is highlighted in bold. RNA aptamers contain 3 extra G residues at their 5' end (written in italics) because they were originally obtained by *in vitro* transcription using the primer M2-F, which contains a G trinucleotide downstream the T7 RNA polymerase promoter.

Aptamer Name	Sequence (5'→3')	Length(nt)
AptD-421	GCGGATCCAGACTGGTGT**ATCTTTTGCATGACGACTATTTACTAACTTTTCTCTAG**GCCCTAAAGACAAGCTTC	74
AptD-1312	GCGGATCCAGACTGGTGT**GCCGTATCCCTCCCTTGTAATTATTTGTTCCATCCGTTCC**GCCCTAAAGACAAGCTTC	76
AptD-1932	GCGGATCCAGACTGGTGT**GGTGGTAATTTTGTGTCTGGTTTGTTCTCTTTCTCGTTGT**GCCCTAAAGACAAGCTTC	76
AptD-2251	GCGGATCCAGACTGGTGT**GTGCGGGTGGGTGGTTTGTTCTGCTGTGTTGTTGTG**GCCCTAAAGACAAGCTTC	76
AptR-421	*GGG*GCGGAUCCAGACUGGUGU**AUCUUUUGCAUGACGACUAUUUACUAACUUUUCUCUAG**GCCCUAAAGACAAGCUUC	77
AptR-1312	*GGG*GCGGAUCCAGACUGGUGU**GCCGUAUCCCUCCCUUGUAAUUAUUUGUUCCAUCCGUUCC**GCCCUAAAGACAAGCUUC	79
AptR-1932	*GGG*GCGGAUCCAGACUGGUGU**GGUGGUAAUUUUGUGUCUGGUUUGUUCUCUUUCUCGUUGU**GCCCUAAAGACAAGCUUC	79
AptR-2251	*GGG*GCGGAUCCAGACUGGUGU**GUGCGGGUGGGUGGUUUGUUCUGCUGUGUUGUUGUG**GCCCUAAAGACAAGCUUC	79

high number of sequences present in the final round of RNA selections, which suggests a more flexible HCV core-binding pattern. Nevertheless, a detailed inspection of the common 6-mer motifs in RNA networks shows the ubiquity of variable GU-rich motifs among the selected sequences (such as GUUGUU, GUUUGU, UUGUUG, UUGUUU, GUGUGU, UGUGUU, UGUUGU, UUGUGG, UUUGUU) or the G-only motif GGGGGG. This fact, together with the large number of different sequences present in the final pools of RNA selections, points towards a moderate binding preference of HCV core protein to GU-rich RNA motifs.

The analysis of the most frequent motifs predicted by HOMER software (see Materials and Methods) in each of the selected aptamers (Supplementary Table S6) showed similar results. Given this list of motifs, those that were also found in any of the HCV complete genomes of genotypes 1 to 4 stored in Los Alamos database (LANL) were filtered out. The most frequent motifs present in AptD-1932 (mainly in DNA selections) and AptD-2251 (in both DNA and RNA selections) are variations or extensions of the tTTGTTc motif (such as TTGTTCCA, TTGTTCCATCGT, TTTTYTYC, TTYYTTC and TTTGYT/ARCAAA; Y and R stand for C/T and A/G, respectively) and are present in all HCV genomes analyzed, with an average of occurrences ranging from 30 to 80 times per genome. Motifs frequent in AptD-421 were found in both RNA and DNA selections, though their sequences are quite different from those present in AptD-1932 and AptD-2251. Of relevance, the most frequent motifs found in AptD-1312 are present in the remaining highly represented aptamers.

Quantification of Kd values using ELONA-based methodologies

The four highly represented aptamer sequences AptD-1312, AptD-421, AptD-1932 and AptD-2251 (both in DNA and RNA versions, see Table 2) were chemically synthesized and assayed using either colorimetric ELONA (DNA aptamers) or ELONA-RTqPCR (RNA aptamers) to quantify their *Kd* values against different HCV core variants, as summarized in Table 3. Negative control molecules D-ACTG and R-ACUG (see Materials and Methods) were tested in parallel (Supplementary Figure S4).

The DNA aptamers showed high binding affinity for the different variants of core protein assayed by colorimetric, indirect ELONA (Figure 3(A)), with *Kd* values in the nanomolar or even high picomolar range (Table 3(A)). The highest affinity was achieved against the variant G1-122, as shown in Figure 4(A), with *Kd* values of 0.4 nM (aptamer AptD-2251), 0.5 nM (AptD-1312), 1.2 nM (AptD-1932) and 20 nM (AptD-421). Therefore, the first three aptamers behave as the best

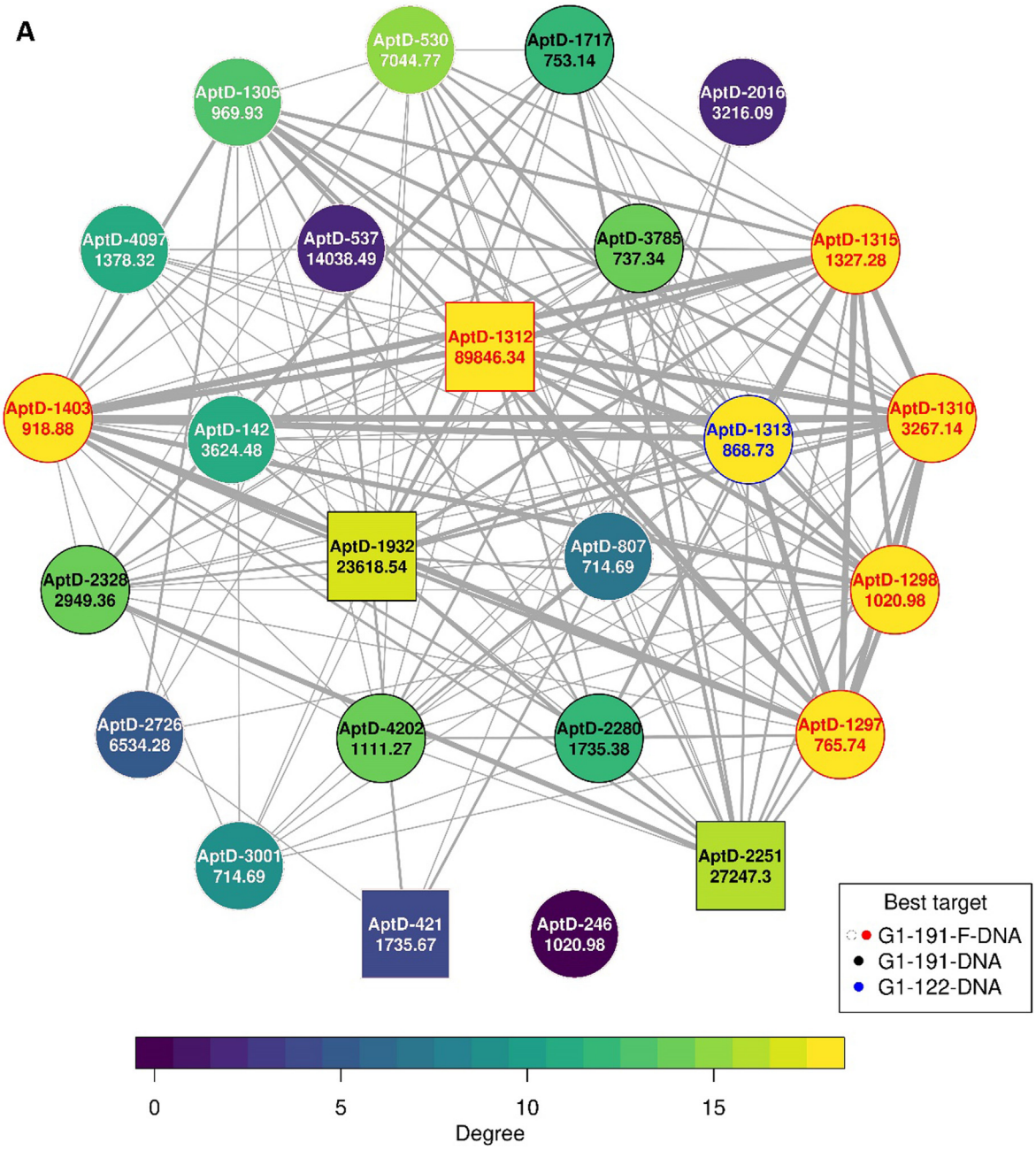


Figure 2. Network of the most represented aptamers found in the last round of DNA (**A**) and RNA (**B**) selection processes against the 6 HCV core targets used. Two aptamers are connected if their sequences share at least 1 out of the 5 most frequent 6-mer motifs present in their variable, central region plus two flanking nucleotides at each side. The weight of the network edges (i.e., the thickness of the grey lines connecting pairs of nodes) is proportional to the number of shared 6-mer motifs. The highly represented aptamers AptD-1312, AptD-421, AptD-1932 and AptD-2251 are identified with squares. Color of the node (from purple to yellow) shows the degree of each aptamer in the network, which is the number of aptamers connected to it. Text color in the aptamer names corresponds to its best HCV core target (see inserts), i.e., the HCV core protein variant for which each aptamer showed the highest CPM value (Supplementary Table S4).

binders for the hydrophilic domain of HCV core of genotype 1 in native conditions, regardless of having been mainly selected against it (AptD-1312) or against the denatured, full-length protein variants G1-191 (AptD-2251 and AptD-1932) or G1-191-F (also AptD-1312, see Figure 1 and Supplementary Table S4).

Due to the fact that AptD-1312 had been selected in all DNA SELEX processes performed, its *Kd* was quantified against the six HCV core variants. This

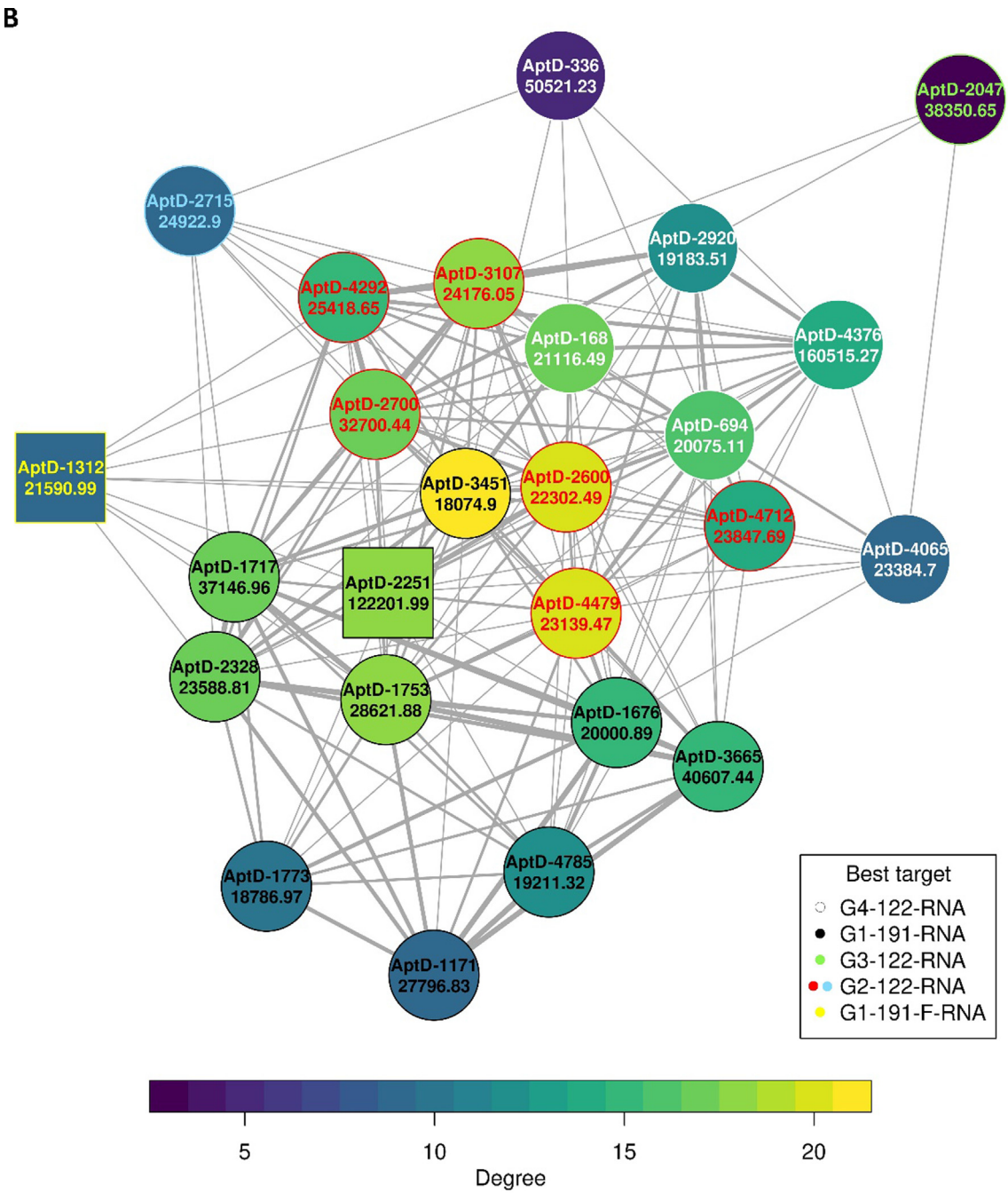


Fig 2. *(continued)*

aptamer showed very high affinity for the hydrophilic domains of the four genotypes, with *Kd* values of 0.5, 2.4, 2.2 and 1.1 nM (G1-122, G2-122, G3-122 and G4-122, respectively). In turn, *Kd* values of 20 and 207 nM were obtained using the full-length proteins G1-191 and G1-191-F, respectively (Table 3(A) and Figure 4(B)). The specificity of AptD-1312 for HCV core was assayed using the proteins bovine serum albumin (BSA), human serum albumin (HSA, which is the most abundant protein in human blood plasma) and the HCV nonstructural protein 3 (NS3) as alternative targets in colorimetric ELONA, in parallel to the four variants of the hydrophilic domain of HCV core (Supplementary Figure S5). Additionally, the other three highly represented DNA aptamers were assayed against G1-122, HSA and NS3 in parallel (Supplementary Figure S6(A)). These experiments revealed a residual binding of the four DNA aptamers to non-specific human and viral proteins, showing statistically significant differences in binding affinities compared to those derived from the specific interaction with G1-122. Among them, AptD-1312 showed the highest

Table 3 ELONA-based affinity constants *Kd* (nM) quantified for the four highly represented DNA (**A**) and RNA (**B**) aptamers selected against the six variants of HCV core protein. The best fit curve corresponded to a one site-specific binding model (with R^2 values of 0.99) in all cases. Each data point represents the mean ± SD (n = 6).

A			B		
Target	**DNA aptamer**	***Kd* ± SD (nM)**	**Target**	**RNA aptamer**	***Kd* ± SD (nM)**
G1-191-F	AptD-421	197 ± 18	**G1-191-F**	AptR-1312	91 ± 9
	AptD-1312	207 ± 20	**G1-191**	AptR-1312	123 ± 11
G1-191	AptD-1312	20 ± 2		AptR-421	1,089 ± 92
	AptD-1932	158 ± 15		AptR-2251	148 ± 13
	AptD-2251	18 ± 2		AptR-1932	47 ± 4
G1-122	AptD-1312	0.50 ± 0.03	**G1-122**	AptR-1312	57 ± 3
	AptD-421	20 ± 1		AptR-421	410 ± 42
	AptD-1932	1.2 ± 0.1		AptR-2251	212 ± 24
	AptD-2251	0.40 ± 0.03		AptR-1932	2,086 ± 77
G2-122	AptD-1312	2.4 ± 0.2	**G2-122**	AptR-1312	41 ± 4
G3-122	AptD-1312	2.2 ± 0.1	**G3-122**	AptR-1312	19 ± 2
G4-122	AptD-1312	1.1 ± 0.1	**G4-122**	AptR-1312	38 ± 3

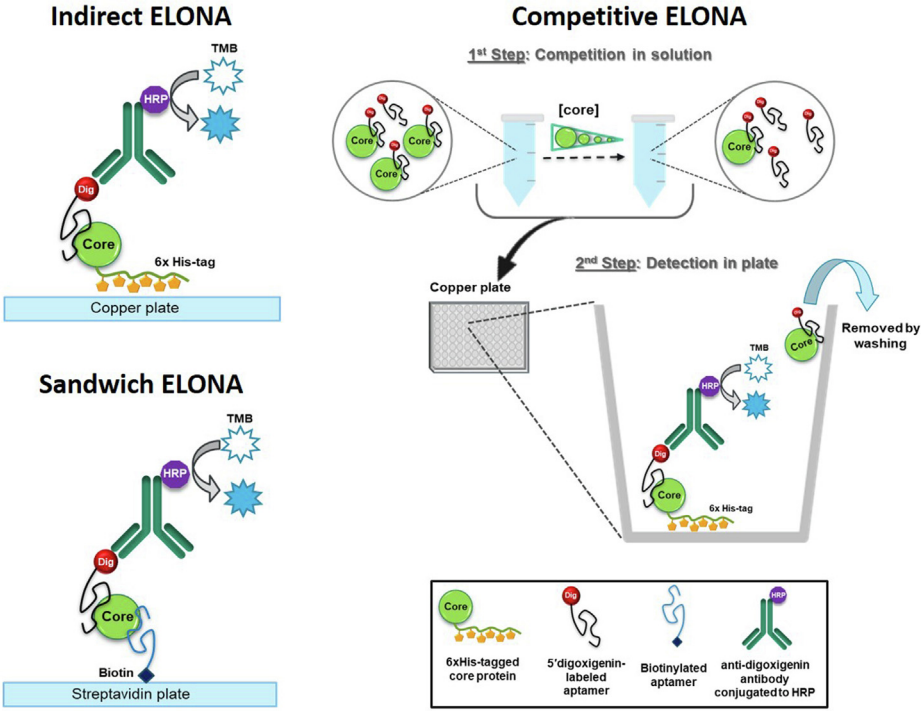


Figure 3. Schematic representation of different colorimetric ELONA formats used in this work: indirect ELONA (**A**), sandwich ELONA (**B**) and competitive ELONA (**C**).

specificity for HCV core, followed by AptD-1932. Of the two proteins used for this specificity control, the higher binding capacity of HSA to the aptamers tested may be due to the possibility that it binds with low affinity to single-stranded nucleic acids, as recently described.[50]

The RNA version of the four highly represented aptamers were assayed using ELONA-RTqPCR, based on a previously described methodology.[49] The *Kd* values obtained using AptR-1312 against the 6 variants of HCV core laid in the low to mid nanomolar range: 91, 123, 57, 41, 19 and 38 nM (for G1-191-F, G1-191, G1-122, G2-122, G3-122 and G4-122, respectively, Figure 5(A)). In turn, the *Kd* of AptR-2251 for the full-length HCV core protein G1-191 (in whose selection process, 12RS10, was the most abundant aptamer found, see Supplementary Table S4 and Figure 1) was 148 nM, while those of AptR-421 and AptR-1932 were 1,089 and 47 nM, respectively (Table 3(B) and Figure 5(B)). Additionally, their affinity for G1-122 protein was assayed and the following *Kd* values were obtained:

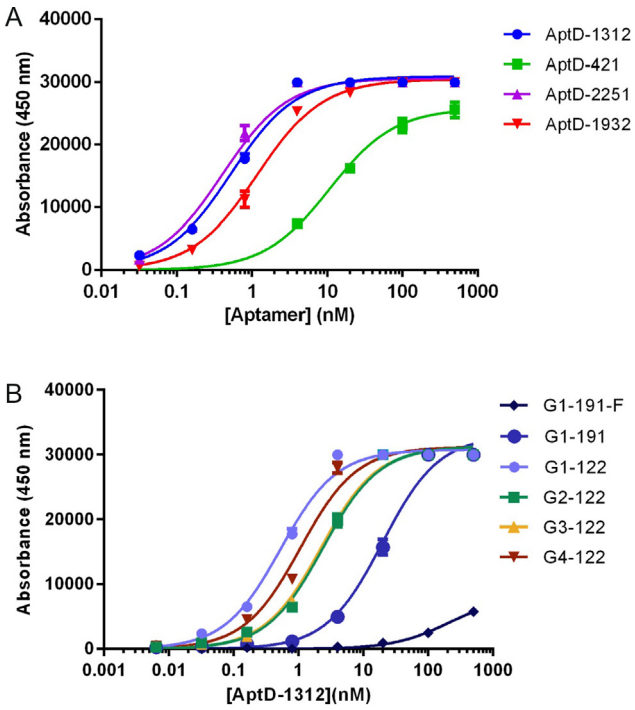


Figure 4. (**A**) Characterization of the binding affinities of the four highly represented DNA aptamers for HCV core G1-122, by means of colorimetric ELONA. The affinity data showed that the best fit curve corresponded to a one site-specific binding model (with R^2 values of 0.99) in all cases, from which the *Kd* ± SD was derived: 0.40 ± 0.03 nM (for AptD-2251), 0.50 ± 0.03 nM (AptD-1312), 1.2 ± 0.1 nM (AptD-1932) and 20 ± 1 nM (AptD-421). Each data point represents the mean ± SD (n = 6). (**B**) Characterization of the binding affinities of AptD-1312 for the six variants of HCV core, using colorimetric ELONA. The affinity data showed that the best fit curve corresponded to a one site-specific binding model (with R^2 values of 0.99) in all cases, from which the *Kd* ± SD was derived: 0.50 ± 0.03 nM (G1-122), 2.4 ± 0.2 nM (G2-122), 2.2 ± 0.1 nM (G3-122), 1.1 ± 0.1 nM (G4-122), 20 ± 2 nM (G1-191) and 207 ± 20 nM (G1-191-F). Each data point represents the mean ± SD (n = 6).

AptR-2251, 212 nM; AptR-421, 410 nM; AptR-1932, 2,086 nM (Table 3(B) and Figure 5(C)). The specificity controls of the four RNA aptamers showed statistically significant differences among their affinity against G1-122 and those against HSA and NS3 proteins at different aptamer concentrations (Supplementary Figure S6(B)). Among them, AptR-1932 showed the highest specificity at all aptamer concentrations tested. A comparative analysis evidenced that the specificity for HCV core protein was higher in DNA aptamers than in RNA ones.

Additionally, 22 other DNA or RNA aptamers that showed high CPM in the last round of at least one of the SELEX processes (Supplementary Table S4) and/or contained sequence motifs shared with the four highly represented aptamers (Supplementary Table S6) were chemically synthesized and analyzed using ELONA-based methods. Three high affinity DNA aptamers were identified for G1-122, with *Kd* values of 2.29, 4.90 and 6.53 nM (AptD-2280, AptD-1399 and AptD-3783, respectively), while the remaining aptamers assayed showed *Kd* values in the low to mid nanomolar range (Supplementary Table S5).

Biosensing potential of DNA aptamers in sandwich and competitive assays

Sandwich and competitive ELONA assays were developed to analyze the usefulness of the high affinity aptamers for the detection of HCV core protein G1-122 in solution. In the sandwich assay (Figure 3(B)), biotin-labelled AptD-1312 was used as capture aptamer and, upon its binding to G1-122 protein, each of the four digoxigenin-labelled aptamers (AptD-421, AptD-1312, AptD-1932 and AptD-2251) were added as detecting aptamers. The four detecting aptamers gave positive results (including AptD-1312, thus suggesting the possible existence of alternative binding sites for it in the hydrophilic domain of HCV core of genotype 1) in a range of G1-122 concentrations between 4 and 64 nM (Figure 6(A)). The calculated limit of detection (LoD) of the assay was as low as 0.002, 0.066, 0.095 and 0.140 nM using AptD-1312, AptD-1932, AptD-421 and AptD-2251 as detecting aptamers, respectively.

For the competitive ELONA assay (Figure 3(C)), digoxigenin-labelled AptD-1312 was incubated with different concentrations of G1-122 core in solution (from 0.02 to 250 nM) and, later on, the

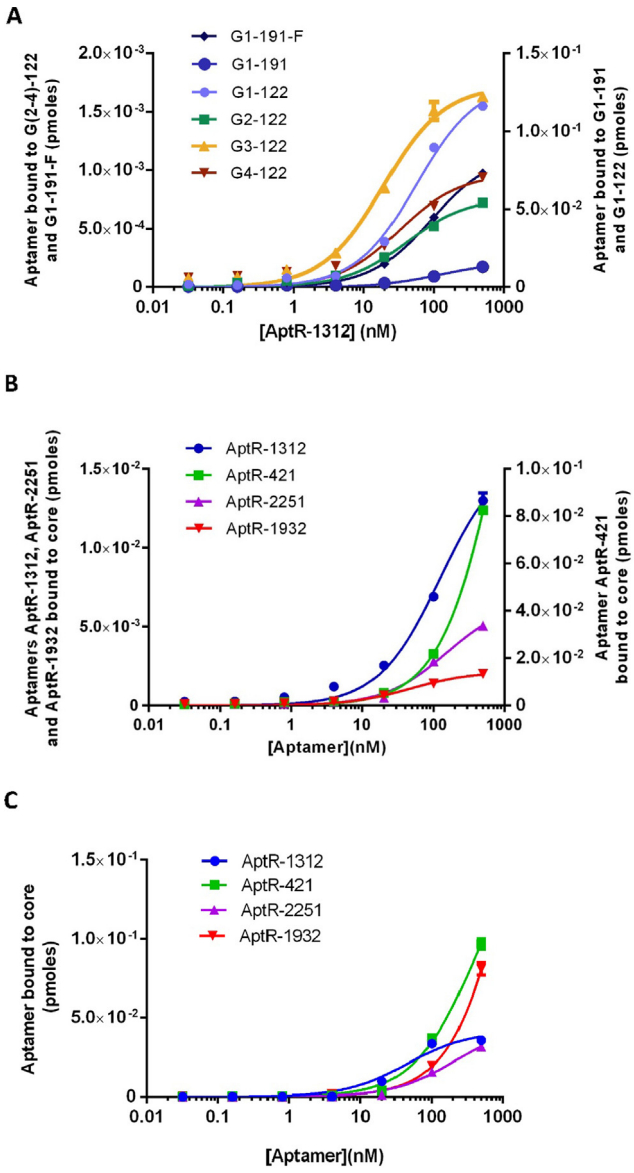


Figure 5. (**A**) Characterization of the binding affinities of AptR-1312 for the six variants of HCV core, using ELONA-RTqPCR (values corresponding to targets G1-122 and G1-191 are shown on the right Y axes). The best fit curve corresponded to a one site-specific binding model (with R^2 values of 0.99) in all cases, from which the *Kd* ± SD was derived: 91 ± 9 nM (G1-191-F), 123 ± 11 nM (G1-191), 57 ± 3 nM (G1-122), 41 ± 4 nM (G2-122), 19 ± 2 nM (G3-122) and 38 ± 3 nM (G4-122). Each data point represents the mean ± SD (n = 6). (**B**) Binding affinities of AptR-1312, AptR-2251 and AptR-1932 (left Y axes), as well as AptR-421 (right Y axes) for G1-191, using ELONA-RTqPCR. The best fit curve was a one site-specific binding model (with R^2 values of 0.99) in all cases, from which the *Kd* ± SD was derived: 1,089 ± 92 nM (AptR-421), 123 ± 11 nM (AptR-1312), 47 ± 4 nM (AptR-1932) and 148 ± 13 nM (AptR-2251). Each data point represents the mean ± SD (n = 6). (**C**) Binding affinities of the four RNA aptamers for G1-122, using ELONA-RTqPCR. The best fit curve was a one site-specific binding model (with R^2 values of 0.99) in all cases, from which the *Kd* ± SD was derived: 410 ± 42 nM (AptR-421), 57 ± 3 nM (AptR-1312), 2,086 ± 77 nM (AptR-1932) and 212 ± 24 nM (AptR-2251). Each data point represents the mean ± SD (n = 6).

unbound aptamer was allowed to react with G1-122 core immobilized onto multiwell plates. The colorimetric results showed that, as expected, higher concentrations of G1-122 in solution left less aptamer available to bind to the immobilized core protein, thus lowering the measured absorbance signal (Figure 6(B)). The calculated LoD of this assay was 2.5 nM.

Finally, the competition among different aptamers for the same (or overlapping) binding site(s) in HCV core G1-122 was assayed in solution. The results showed that digoxigenin-labelled AptD-1312

competed with the other three high affinity, non-labelled aptamers in the order: AptD-2251 > AptD-421 ≫ AptD-1932 (Figure 6(C)). The residual competition between AptD-1312 and AptD-1932 is coherent with their good performance as the pair of aptamers used in sandwich ELONA (Figure 6 (A)).

Effect of aptamers AptD-1312 and AptD-1932 on HCV progeny, extracellular and intracellular viral RNA production

Highly represented and high affinity aptamers AptD-1312 and AptD-1932, which did not compete for the same binding sites on HCV core G-122 (Figures 6(A) and (C)), were studied in cell culture for their possible effect on viral progeny, extracellular and intracellular viral RNA production. With that aim, Huh-7.5 cells were either mock-transfected or transfected with aptamers AptD-1312 or AptD-1932. Cells were transferred to a larger culture plate 24 h post-transfection; 24 h later, cells were infected with HCV p0[51] at a multiplicity of infection (MOI) of 0.1 $TCID_{50}$/cell. Viral titer, and extracellular and intracellular RNA levels were measured 72 h post-infection. Parallel transfections were carried out using the ssDNA molecule D-ACTG as a negative control (Figure 7(A)).

The use of fluorescein-labelled aptamers showed that they accumulate in Huh-7.5 cells (Supplementary Figure S7). The results of the functional assays indicated that the decrease of viral progeny titer produced by aptamers AptD-1312 and AptD-1932 was 1.9-fold ($p = 0.002$; t-test) and 3.4-fold ($p = 0.0015$; t-test), respectively (Figure 7(B)). The decrease in extracellular viral RNA level was 8.9-fold ($p = 0.024$; t-test) for AptD-1312 and 13.8-fold ($p = 0.021$; t-test) for AptD-1932 (Figure 7(C)). Concomitantly, the intracellular viral RNA increased 4.9-fold ($p = 0.0045$; t-test) for AptD-1312 and 5.2-fold ($p = 0.00076$; t-test) for AptD-1932 (Figure 7(D)). Thus, the results confirmed the inhibitory effect of aptamers AptD-1312 and AptD-1932 on HCV p0 infectious viral progeny and extracellular RNA production.

Discussion

Aptamers are increasingly used in virology, including those selected against HCV targets (reviewed in[42,45]). RNA and DNA aptamers specific for the structured 5' (containing the IRES element) or 3' UTR of HCV genomic RNA have been reported.[52–55] Also, different aptamers have been developed against non-structural and structural HCV proteins, including the viral proteases NS2[56] and NS3,[57–59] the dimeric metalloprotein NS5A (which is an essential component of the HCV replication complex),[60] the RNA-dependent RNA polymerase NS5B,[61–63] the envelope proteins E1 and E2[64–66] and core protein.[46–48] However, despite being a highly conserved and multifunctional protein, HCV core has still been little explored as the target for aptamer-based viral diagnosis and for the design of therapeutic strategies alternative or complementary to DAAs.

The first HCV core-specific aptamers were selected against the N-terminal amino acids 2 to 114 of HCV core of genotype 1 expressed in *E. coli*,[46] in a process that yielded four RNA aptamers showing nanomolar affinity for its target (*Kd* values of 142 and 224 nM). The authors developed a sandwich assay using labelled antibodies as reporters, though no LoD or sensitivity of the method was reported. Using one of these aptamers, a lateral flow assay for the direct detection of HCV core antigen was further developed, with a LoD of 10 pg/ml, which corresponded with a viral load of 10^4 copies/ml.[67] Later on, other authors reported the selection of 7 DNA aptamers against the full-length HCV core protein of genotype 2 expressed in *E. coli*.[47] Although their *Kd* values were not quantified, some of them were shown to inhibit the production of infectious virus in cell culture (measured as the secretion of core protein into the culture medium of virus-infected cells) and were used to detect HCV in serum samples from patients infected with various genotypes of the virus. As a result, different groups have developed HCV sensing platforms based on some of the DNA aptamers selected by Shi et al. in 2014,[68–69] whose improved sensitivity is promising for future use in clinical practice.

In this work, we have selected DNA and RNA aptamers in parallel against six variants of HCV core protein, through 12 independent SELEX processes. Three different variants of genotype 1 HCV core protein have been used: two full-length versions (G1-191, and the β-galactosidase-tagged fusion protein G1-191-F) and the complete N-terminal hydrophilic domain (G1-122). Additionally, the hydrophilic domains of HCV core protein belonging to the genotypes 2, 3 and 4 (G2-122, G3-122 and G4-122) have also been used as selection targets. Another variable explored was the fact that G1-191 and G1-191-F were available in denatured form, while the four 122 aa-long versions of the protein were used in native state. Thus, we designed a multivariate strategy with the aim of obtaining the best and most universal aptamer candidates for HCV diagnosis and inhibition. Notably, five out of the six proteins used as selection targets were expressed in baculovirus-insect systems instead of in *E. coli*, and thus are assumed to contain post-translational modifications similar to those present in HCV-infected human cells.

The bioinformatics analysis of the highly represented sequences obtained in the last round of each DNA and RNA SELEX process showed a

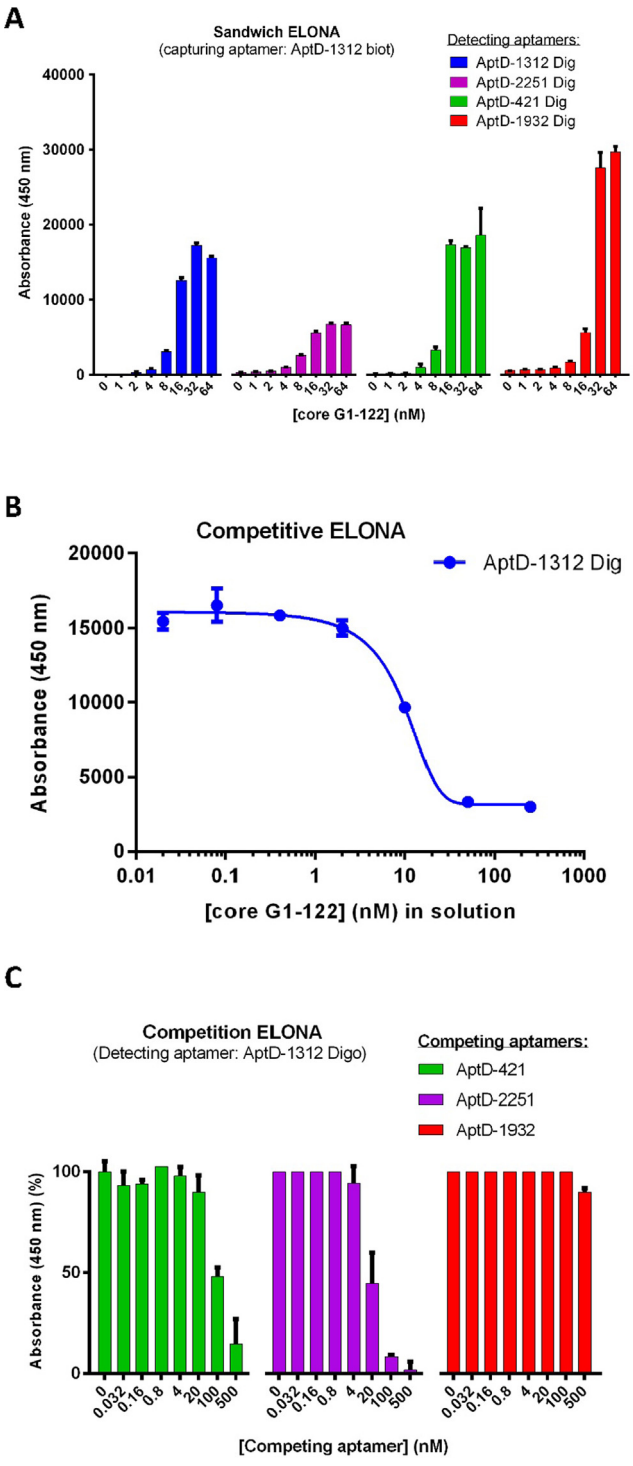


Figure 6. Biosensing potential of the selected DNA aptamers. (**A**) Colorimetric Sandwich ELONA using the biotin-labelled AptD-1312 as capture aptamer for G-122 core protein, and the digoxigenin-labelled detecting aptamers AptD-1312, AptD-2251, AptD-421 and AptD-1932. Each data point represents the mean ± SD (n = 3). (**B**) Competitive ELONA using AptD-1312 aptamer and HCV core protein G1-122 in solution (ranging from 0 to 250 nM). Each data point represents the mean ± SD (n = 3). The obtained R^2 value was 0.99 (**C**) Binding site competition assay in solution to evaluate the possibility that aptamers share the same (or overlapping) binding site(s) in G1-122 core protein. In this ELONA assay, non-labelled or "cold" aptamers (AptD-421, AptD-2251 and AptD-1932) were used as competing aptamers and digoxigenin-labelled AptD-1312 was the detecting aptamer. Each data point represents the mean ± SD (n = 3).

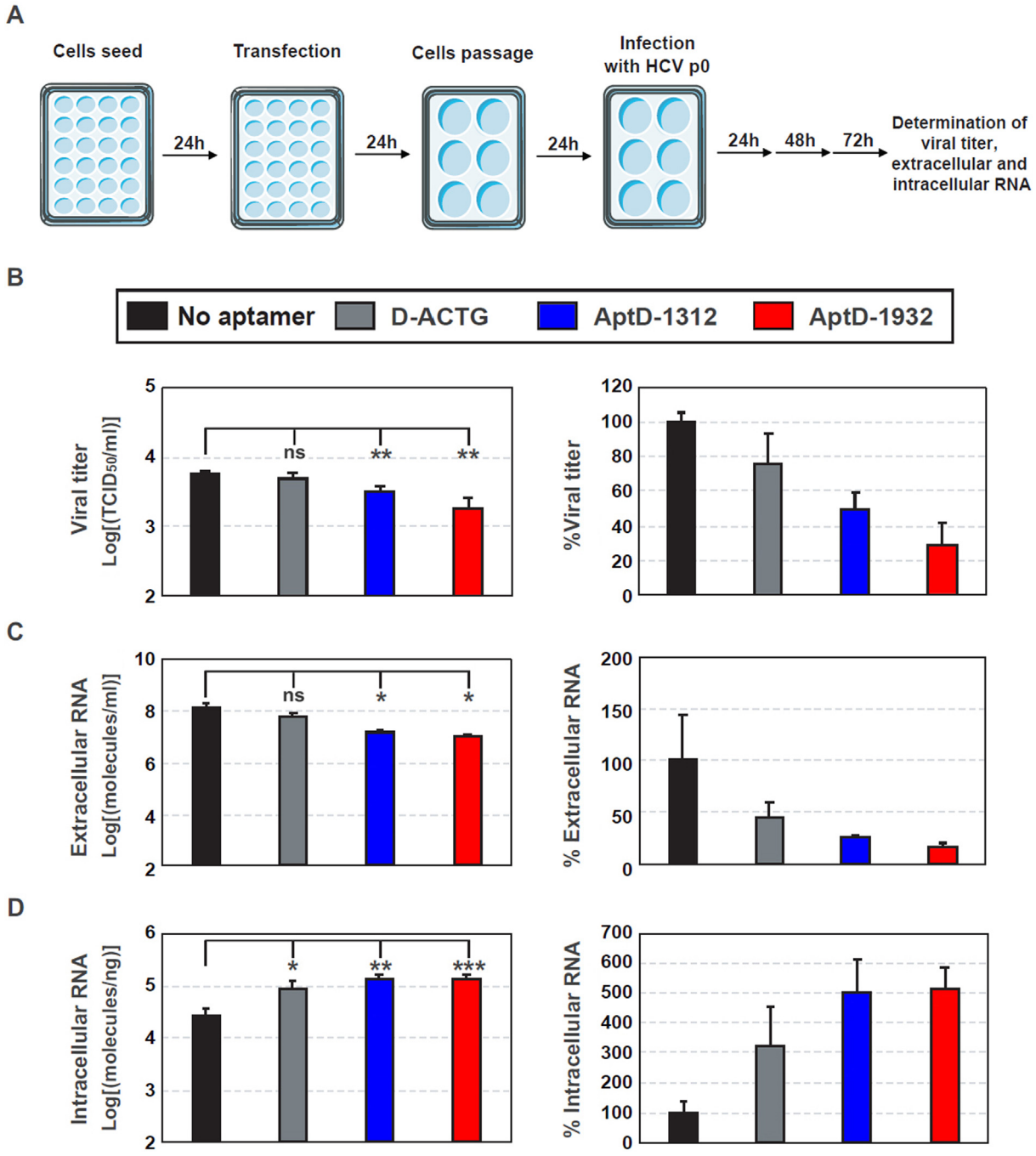


Figure 7. Effect of DNA aptamers AptD-1312 and AptD-1932 on HCV progeny, extracellular and intracellular viral RNA production. (**A**) Experimental scheme. Huh 7.5 cells were seeded in 24-well plates, and 24 h later cells were transfected with the aptamers. Twenty-four hours after transfection, cells were transferred to 6-well plates and infected with HCV p0 24 h later. Supernatants and intracellular RNA were collected 72 h post-infection. (**B**) Effect of the aptamers on infectious progeny production of HCV p0. The aptamer used is indicated in the upper box. The viral titer expressed as $\log_{10}$ $TCID_{50}$ /ml is indicated in the ordinate. Values of viral titer (mean ± SD, n = 3) are $6.08 \times 10^3 \pm 3.98 \times 10^2$ (HCVp0), $4.81 \times 10^3 \pm 1.09 \times 10^3$ (D-ACTG), $3.11 \times 10^3 \pm 5.97 \times 10^2$ (AptD-1312) and $1.79 \times 10^3 \pm 8.75 \times 10^2$ (AptD-1932). The graphic on the right side represents the viral titer expressed as a percentage. (**C**) Same as B, but for the extracellular viral RNA (mean ± SD, n = 3): $1.47 \times 10^8 \pm 6.40 \times 10^7$ (HCVp0), $6.56 \times 10^7 \pm 2.21 \times 10^7$ (D-ACTG), $1.64 \times 10^7 \pm 1.94 \times 10^6$ (AptD-1312) and $1.07 \times 10^7 \pm 1.94 \times 10^6$ (AptD-1932). (**D**) Same as B, but for the intracellular viral RNA (mean ± SD, n = 3): $2.73 \times 10^4 \pm 9.81 \times 10^3$ (HCVp0), $8.81 \times 10^4 \pm 3.53 \times 10^4$ (D-ACTG), $1.36 \times 10^5 \pm 3.13 \times 10^4$ (AptD-1312) and $1.41 \times 10^5 \pm 1.90 \times 10^4$ (AptD-1932). All titrations were carried out in triplicate and the statistical significance of the differences was determined by t-test: ns, no significant difference; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

convergence towards a limited number of aptamers. Among them, AptD-1312, AptD-421, AptD-1932 and AptD-2251 were chosen for further ELONA-based biochemical analysis in both DNA and RNA versions. They were clearly different in sequence and secondary structure (both globally and after the motif analysis performed) from the HCV core-specific aptamers previously published by other authors.[46–48] The only exception is that the sequence alignment of AptD-1932 and the aptamer Cnew published by Shi et al.[47] shows an overall nucleotide identity of 46.4%, which is distributed in sequence stretches not longer than 5 nt and is not reflected at all in their respective secondary structures. Additionally, a very limited number of the short sequence motifs found in the variable regions of some of the highly prevalent aptamers selected in this work (in particular, KCCCTc and tTTGTTc) are also present in aptamers C7, C42, C104 and Cnew published by Shi et al.[47]. Regarding the RNA aptamers, the lack of similarity between those obtained in this work and those previously published by Lee et al.[46] is consistent with the high variability observed in the final rounds of the RNA selection processes carried out here. The relevance of these findings will be investigated in a future work.

The aptamer sequence AptD-1312 was retrieved in the 12 SELEX processes performed, being highly represented in the 6 DNA selections. This fact suggests that the target sequence of such a universal or "pangenotypic" aptamer could be present inside the hydrophilic domain of core protein, in a region which seems conserved among genotypes 1 to 4 and accessible in both native and denatured states of the protein. However, the lower *Kd* values of AptD-1312 for the hydrophilic domains (0.5, 2.4, 2.2 and 1.1 nM for core variants G1-122, G2-122, G3-122 and G4-122, respectively) than for the complete protein (20 and 207 nM for G1-191 and G1-191-F, respectively) points towards a tighter aptamer-protein recognition when the polypeptide is accessible in native conditions. This fact also suggests that the fusion protein G1-191-F could exhibit some steric hindrance towards the binding of AptD-1312. The bioinformatics characterization of short sequence motifs present in the selected aptamers suggests that HCV core protein may bind to different nucleic acid sequences, including variants of the tTTGTtc and TCCCT motifs as well as other GT-rich sequences. Different versions of these motifs are present in AptD-1312 and remain accessible in its predicted secondary DNA structure. This could be one of the reasons why this aptamer efficiently recognizes every HCV core variant assayed, albeit with different affinity. In turn, the RNA aptamer AptR-1312 showed higher *Kd* values against the 6 variants of HCV core, lying in the low to mid nanomolar range.

The aptamer AptD-2251 showed a *Kd* of 0.4 nM for G1-122, which is even smaller than that of AptD-1312. Therefore, although the methodologies used to quantify *Kd* values are different, the affinity of these two (DNA) aptamers is 280 to 560 times higher than that of the only two (RNA) HCV core-specific aptamers for which *Kd* values were previously reported: 142 nM for the aptamer termed "9–14" and 224 nM for the aptamer "9–15" published by Lee et al.[46]. Regarding the RNA aptamers selected in this work against G1-122, the affinity shown by AptR-1312 (with *Kd* of 57 nM) is 2.5 to 6.4 times higher than that of the RNA aptamers selected by Lee et al. against a similar target (containing aa 2–114 of the hydrophilic domain of HCV genotype 1).[46]

Most of the *Kd* values of the four highly represented aptamers laid in the picomolar to nanomolar range for different HCV core targets. Additionally, all of them displayed high binding specificity for HCV core compared to that shown against HSA (the most abundant protein in human blood plasma) and HCV NS3. Therefore, these high-affinity and high-specificity aptamers provided various options for analyzing their pairwise behavior, which is required in sandwich and competitive ELONA-based biosensing. The calculated LoD of the sandwich assay (with AptD-1312 as the capturing aptamer and HCV core G1-122 as the target protein) was 0.002, 0.066, 0.095 and 0.140 nM using AptD-1312, AptD-1932, AptD-421 and AptD-2251 as detecting aptamers, respectively. In these assays, the fact that the binding of AptD-1312 to G1-122 does not avoid the further binding of other AptD-1312 molecule to the same target suggests that the hydrophilic domain of HCV core of genotype 1 might expose at least two alternative binding sites for this aptamer. However, since G1-122 does not contain repeated sequence domains, both potential binding sites for AptD-1312 are expected to show different binding affinities. This possibility will be investigated in a future work. Regarding competitive ELONA assays, a calibration curve between 0.016 and 250 nM of HCV core was constructed, from which a LoD of 2.5 nM was calculated. The very high sensitivity achieved using these aptamer-based assays paves the way to the development of novel and highly efficient diagnostic systems for HCV, with applicability as POC tests.

The results of sandwich and competitive ELONA showed that, among the highly represented DNA aptamers that exhibit high affinity and specificity for HCV core, AptD-1312 and AptD-1932 were the least likely to compete with each other for their binding sites in the hydrophilic domain of this protein. Therefore, both aptamers were assayed in independent cell culture experiments (following previous reports[47]) as they could provide alternative pathways to inhibit HCV production based on their interaction with the essential and multifunctional core protein. They were shown to decrease the viral progeny titer and the extracellular viral RNA level,

with a concomitant increase in intracellular viral RNA. These results suggest that AptD-1312 and AptD-1932 inhibit HCV capsid assembly and, thus, virion formation and exit in cell culture.

As a consequence, the aptamer-driven intracellular accumulation of HCV core protein not assembled in capsids is expected, which might entail an increase in associated lipid droplets. Additional experiments are needed to explore these possibilities and will be conducted in a future work. Moreover, since Huh-7.5 cells are known to be deficient in innate immune signaling,[70] possible additional effects of these aptamers on immune activation cannot be ruled out.

Future experiments will also include the biochemical analysis of other highly represented aptamers selected in the 12 SELEX processes performed, and the optimization of some of the aptamers that show high affinity and specificity to get minimal HCV core-binding sequences. The best DNA aptamers will be tested in ELONA assays and in different biosensor formats using HCV positive sera from infected patients (as well as uninfected control sera), which will allow us to compare their performance with that of standard diagnostic tools used in the clinical practice. In parallel, given the already demonstrated antiviral activity of AptD-1312 and AptD-1932, the inhibitory potential of other high affinity and highly represented DNA and RNA aptamers (in the latter case, once the required chemical modifications are introduced to increase their resistance to degradation by cellular RNases) will be tested individually and in combination, and the delivery mechanisms will be optimized. This will hopefully contribute to the design of novel aptamer-based strategies for the treatment of hepatitis C.

Materials and methods

Target HCV core proteins

Six variants of HCV core protein were used as targets for the selection of DNA and RNA aptamers in this work (Supplementary Figure S1). Two of them contained the full-length sequence of HCV core (aa 1–191), belonging to genotype 1 and available in denaturing conditions: a commercial (Abcam, Cat. No. ab49020) C-terminal biotin-tagged, fusion protein with β-galactosidase expressed in *E. coli* that we have termed G1-191-F, and an N-terminal 6xHis-tagged protein (see below) expressed in baculovirus-insect systems (Algenex, Spain)[71] termed G1-191. The four remaining variants (also expressed in baculovirus-insect systems by Algenex) contained the 6xHis-tagged, hydrophilic domain (aa 1–122) of HCV core in native conditions belonging to genotypes 1 to 4, termed G1-122, G2-122, G3-122 and G4-122, respectively.

Representative sequences of HCV core protein belonging to genotypes 1 to 4 were previously chosen for being produced in baculovirus-insect systems. First, a BlastP search[72] of HCV core proteins available in GenBank was performed, using a seed sequence of known core proteins of each genotype and default parameters (BLOSUM63 matrix, gap open penalty 11, gap extension penalty 1). The sequence alignment was obtained through MAFFT v7.215[73] using options globalpair-maxiterate 1000. The corresponding distance matrix was obtained by means of the program distmat included in EMBOSS v6.6.0.0[74] without correcting for multiple substitutions. The consensus sequence of each alignment was produced using the program cons in EMBOSS v6.6.0.0. For each of the four HCV genotypes, its representative core sequence was chosen as that showing the lowest distance to the corresponding consensus sequence, and thus the minimum sum of distances to all the sequences in the alignment. Their GenBank accession numbers were AFS60333.1, AFS60360.1, ACZ60107.1 and DQ988076 for HCV genotypes 1 to 4, respectively (Supplementary Table S1).

In vitro selection of DNA and RNA aptamers against HCV core proteins

Immobilization of the target proteins for SELEX. To start DNA and RNA aptamer selections against the G1-191-F core variant, a volume of 100 μl (370 pmoles) of the biotin-tagged protein was attached to 100 μl of streptavidin-coated magnetic beads (Dynabeads MyOne Streptavidin T1, Invitrogen) resuspended at 10 mg/ml following manufacturer's instructions. Alternatively, 10 pmoles of the 6xHis-tag proteins (G1-191, G1-122, G2-122, G3-122 or G4-122) were immobilized onto Pierce® Copper Coated High Binding Capacity Plates (Thermo Scientific Fisher), pre-blocked with bovine serum albumin (BSA, Sigma), following the manufacturer's guidelines.

Initial ssDNA library. The ssDNA combinatorial library and the primers used for amplifications were chemically synthesized and HPLC-purified by IBA Lifesciences (Goettingen, Germany). The library, termed M1-40, consisted of a pool of 76 nt-long oligonucleotides containing two flanking primer-binding regions of 18 nt-long each, and a central, 40 nt-long random region: 5'-GCGGATCCAGACTGGTGT(N_{40})GCCCTAAAGACAAGCTTC-3'.

To start the selection of DNA aptamers, 80 pmoles of M1-40 (equivalent to 4.8×10^{13} molecules) were amplified by asymmetric PCR [94 °C 2 min, 25 cycles of (94 °C 15 sec, 60 °C 30 sec, 72 °C 45 sec), 72 °C 5 min]. The amplification was performed using the forward primer M1-F (5'-GCGGATCCAGACTGGTGT-3') (at 0.5 μM) and the reverse, 5'-biotinylated primer

M1-R-Biot (5'-Bio-GAAGCTTGTCTTTAGGGC-3') (0.1 μM), dNTPs (0.2 mM), High Fidelity Expand DNA polymerase (EHF) (0.07 U/μl) (Sigma) and EHF buffer 1X (with 2 mM Mg^{2+}).

The mixture of ssDNA and dsDNA obtained was subjected to strand separation to produce ssDNA using streptavidin magnetic beads.[75] Briefly, 400 μl of MagnaBind Streptavidin beads (Thermo Scientific Fisher) at 5 mg/ml were washed just before use, once with 400 μl of 100 mM NaOH and twice with 600 μl of SB1T Buffer (40 mM HEPES pH 7.5, 125 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.05% Tween-20), and the beads were then resuspended in 400 μl of 1 M NaCl with 0.05% Tween-20. Upon activation of the magnetic beads, 200 μl of amplified, biotinylated dsDNA were added, incubated at 25 °C for 5 min with 14.000 rpm agitation, resuspended and incubated for 5 min in the same conditions. Then, the beads were washed three times using 300 μl of SB1T. Afterwards, 480 μl of 100 mM NaOH were added and the suspension was incubated at 25 °C for 10 min, with 600 rpm stirring. Finally, the sample was partially neutralized by adding 120 μl of 80 mM HCl and 6 μl of 1 M HEPES. The eluted, non-biotinylated ssDNA (forward strand) was purified using Amicon Ultra-0.5 mL 10 K Centrifugal Filters (Merck). Such a purified ssDNA strand was quantified by UV–vis spectrophotometry (Nanodrop, Thermo Fisher Scientific) and its integrity was checked by electrophoresis using a denaturing (5 M urea) 10% polyacrylamide gel.

Alternatively, ssDNA was prepared by means of λ-exonuclease digestion. With that aim, 80 pmoles of M1-40 were PCR amplified using the forward primer M1-F and the 5'-phosphorylated, reverse primer M1-R-5'P (5'-GAAGCTTGTCTTTAGGGC-3'), in the conditions described above. The reverse DNA strand was removed by enzymatic digestion, using 10 units of λ-exonuclease (Thermo Scientific) for every 2 μg of amplified dsDNA, in 50 μl of 1X reaction buffer (67 mM glycine–KOH pH 9.4, 2.5 mM $MgCl_2$, 0.01% Triton X–100). The reaction mixture was incubated at 37 °C for 1 hour (h), followed by heat deactivation at 80 °C for 10 min. The resulting ssDNA (forward strand) was precipitated by adding 0.1 volumes of 3 M NaOAc pH 5.5 and 2.5 volumes of cold absolute ethanol. After precipitation for a minimum of 1 h at −80 °C, the DNA was centrifuged at 4 °C for 40 min at maximum speed. The supernatant was washed using 70% ethanol. After a centrifugation of 15 min, the pellet was dried and resuspended in water. Then, the DNA was purified using Amicon Ultra-0.5 mL 10 K Centrifugal Filters (Merck).

In turn, for the selection of RNA aptamers 80 pmoles of M1-40 were PCR amplified using the forward primer M2-F (5'-*TGTAATACGACTCACT AATA*GGGGCGGATCCAGACTGGTGT-3', which includes the T7 RNA polymerase promoter, in italics) (0.06 μM) and the reverse M1-R (0.06 μM), in a 1 ml reaction, with 4 cycles of the same program described for ssDNA aptamers. Milli-Q water treated with DEPC was used in all the processes involving RNA. The amplified product was *in vitro* transcribed (IVT) by T7 RNA polymerase using the AmpliScribe™ T7 High Yield Transcription Kit (Epicentre) following manufacturer's instructions. The resulting RNA pool was precipitated using 1 volume of 5 M NH_4OAc and 2.5 volumes of cold absolute ethanol, following the protocol described above.

In vitro selection of DNA and RNA aptamers against different variants of HCV core protein. DNA and RNA independent SELEX processes were performed against each of the 6 variants of HCV core protein (Supplementary Figure S1) using up to 14 SELEX rounds, as summarized in Supplementary Table S2. The first round of each DNA aptamer selection was carried out using $\sim 10^{14}$ molecules of the amplified initial ssDNA library (see above). With that aim, the amplified ssDNA population was resuspended in 90 μl of Selection Buffer (SB) (100 mM HEPES pH 7.5, 100 mM NaCl, 10 mM KCl, 5 mM $MgCl_2$, 2 mM $CaCl_2$, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.05% Tween-20), denatured at 95 °C for 10 min and renatured by cooling at 37 °C for 10 min. The folded ssDNA pool was added to either 90 μl of selection matrix containing streptavidin-coated magnetic beads bound to the biotinylated G1-191-F core protein, or to well plates containing the 6xHis-tagged proteins G1-191, G1-122, G2-122, G3-122 or G4-122 (see above). In all cases, the ssDNA pool was incubated with its target at 37 °C for 1 h. In the partition step, unbound ssDNA was removed by washing two times with 500 μl of SB. Then, core-bound aptamers were recovered by adding 100 μL of Milli-Q water, and heating at 95 °C for 10 min. The preparation of the ssDNA population to initiate the next SELEX round was performed by asymmetric PCR amplification of the recovered ssDNA aptamers as described above, using a number of cycles that varied between 20 and 30 along the different rounds and SELEX processes.

To foster the selection of high affinity aptamers, the stringency of the selection process was stepwise increased at each round by doubling the number of washing steps. Also, the incubation time was decreased along all the processes and, in some cases, the amount of ssDNA added to the selection matrix (thus, the ssDNA:protein ratio) was modified, as previously described.[49,76] Counter-selection steps were introduced along the SELEX processes to discard aptamers that could have been selected against the inert matrix (either naked streptavidin beads or empty, streptavidin-coated well plates). Also, counter-selection was

used in most of the processes to increase the specificity of the selected aptamers, by discarding those able to bind to other proteins abundant in human serum (HSA and prothrombin). Additionally, counter-selection was used to discard the aptamers that were able to bind to only one of the HCV core genotypes, since universal or pangenotypic aptamers were sought. In one of the selections, a competition (termed CP) among the selected aptamers and tRNA was also introduced (Supplementary Table S3).

For the *in vitro* selection of RNA aptamers against the six variants of HCV core protein, the first SELEX round was accomplished using $\sim 10^{14}$ molecules of the *in vitro* transcribed RNA pool refolded in 90 μl of SB, following the protocol described above. Once the unbound RNA molecules were discarded, the recovered aptamers in each round were retrotranscribed and amplified by one-step RT-PCR [55 °C 30 min, 94 °C 5 min, (94 °C 15 sec, 60 °C 30 sec, 68 °C 45 sec), 68 °C 5 min] with SuperScript® III reverse transcriptase and Platinum® *Taq* DNA Polymerase (Life Technologies) using M2-F (0.5 μM) and M1-R (0.5 μM) in a total volume of 100 μl. The number of PCR cycles in each SELEX round varied between 20 and 30. After each RT-PCR, 2 μg of the amplified DNA was *in vitro* transcribed and the resulting RNA pool was purified by ethanol precipitation. Other experimental variables were the same than those described for the selection of DNA aptamers.

Ultra-deep sequencing and bioinformatics analysis of the aptamer populations

After the last round of each selection process (Supplementary Table S2), the DNA or RNA pools were re-amplified using 57 nt-long forward standard selection primers containing specific 6 nt-long barcode sequences, as well as a common, 34 nt-long reverse one. Barcoded, 131 nt-long PCR samples were mixed in equimolar amounts and submitted to Sistemas Genómicos, S.L. (Valencia, Spain) for standard UDS-Illumina MiSeq sequencing.

Forward Illumina paired end read sequences were demultiplexed with the AptaPLEX method implemented in the software collection AptaSUITE 0.9.0[77] using the following parameters: total aptamer length between 61 and 81 nt (i.e., length of the variable region between 25 and 45 nt); minimum overlapping region length = 15 nt; maximum mismatches in overlapping region = 5; maximum mismatches in the forward or reverse primer-binding regions = 3. Demultiplexed aptamers were then clustered applying the AptaCluster[78] algorithm, included in the same software, with the default parameters. Seed sequences of clusters with a relative frequency above 100 CPM in at least one of the final SELEX rounds were sorted alphabetically and a code ID was assigned to each cluster after its position. Supplementary Table S4 includes the relative frequency of these sequence clusters in each final selection round.

Over-represented sequence motifs found in the aptamer pools selected in each final round were identified by applying the findMotifs.pl routine, included in the HOMER package,[79] using as background the initial pool of sequences and, when available, the sequences corresponding to a selection step previous to the final one. Two types of motif discovery methodologies were applied: the default search routine for motifs of undetermined length, and a motif-length based approach restricted to small motifs of 5 to 8 nt. Motifs detected by HOMER in the final round of each selection were then compared with the sequence of each one of the over-represented aptamers selected for biochemical validation (see below). They were also compared with the HCV complete genomes of genotypes 1 to 4 stored in Los Alamos HCV Sequence Database,[80] which contained 1,909 genomes of genotypes 1 to 4 (1,571, 186, 71 and 81 respectively). The sequence motifs found in the different selections that are present in each one of the selected high affinity aptamers, as well as their frequency in HCV full genomes is summarized in Supplementary Table S6.

Additionally, short sequence similarity was studied for the 25 most frequent seed sequences that showed at least 50 CPM in each individual protein target, as well as for the combined data of the 6 DNA SELEX processes, the 6 RNA SELEX processes, and the 4 DNA selections against the 122 aa-long targets (G1-122 to G4-122). Short sequence motifs were defined by decomposing the variable, central region, plus two flanking nucleotides at each side, into k-mers by a sliding window of step 1. Then, DNA and RNA aptamer networks were constructed based on short sequence motif co–occurrence of the 5 most frequent motifs.

The secondary structure of the four high affinity DNA and RNA aptamer sequences was predicted using Mfold.[81] For DNA aptamers, the folding parameters were set to values analogous to those operating in the *in vitro* selection cycles (37 °C, 100 mM Na^+ and 5 mM Mg^{2+}). In turn, for RNA aptamers the folding temperature was also 37 °C though the only ionic conditions allowed by the software were 1 M Na^+ and no divalent cations. For comparison, DNA aptamers were also folded in the same ionic conditions than those of RNA. The program QGRS Mapper[82] was used in parallel to detect and predict the formation of G-quadruplexes in DNA and RNA aptamers.

Colorimetric ELONA for the quantification of Kd of DNA aptamers

To biochemically characterize the selected aptamers using ELONA-based methods, 10 pmoles of G1-191-F core variant were

immobilized onto high binding capacity 96-well plates (Corning™ Costar™ 96-Well EIA/RIA Plates, Thermo Scientific Fisher). The plate was incubated overnight at 4 °C, and then at 37 °C for 1 hour with agitation (300 rpm). The supernatant was discarded, 100 μl/well of blocking buffer (PBS 1X with 1% BSA) was added and incubated at 37 °C for 30 min with agitation (300 rpm), and the supernatant was discarded again. In turn, ELONA-based protocols involving the 6xHis-tag proteins (G1-191, G1-122, G2-122, G3-122 or G4-122) were performed on Pierce® Copper Coated High Binding Capacity Plates (Thermo Scientific Fisher) as described above.

An indirect, colorimetric ELONA (Figure 3(A)) was used to quantify the *Kd* values of the most represented DNA aptamers, as previously described.[49] Each individual 5' digoxigenin-labelled ssDNA aptamer (purchased to IBA Lifesciences) was assayed (in triplicate) at eight increasing concentrations (0, 0.032, 0.16, 0.8, 4, 20, 100 and 500 nM). A synthetic, 76 nt-long ssDNA molecule termed D-ACTG was used as a negative control, which contains 10 repeats of the tetranucleotide ACTG in its central region [sequence 5'-G CGGATCCAGACTGGTGT(ACTG)$_{10}$ GCCCTAAAGACAAGCTTC-3']. Two independent multiwell plates were used for each experiment, thus giving rise to 6 aptamer-target data points for each aptamer concentration. The non-linear regression analysis of the binding curves and the *Kd* quantification were performed using GraphPad Prism 6. To test the specificity for HCV core of the highly represented DNA aptamers, colorimetric ELONA assays were conducted following the protocol described above and using three 6xHis-tagged proteins: BSA (abcam), HSA (abcam) and HCV NS3 (RayBiotech).

ELONA-RTqPCR for the quantification of Kd of RNA aptamers

The highly represented RNA aptamers selected were subject to biochemical analysis using ELONA-RTqPCR, as previously described.[49] For that, 10 pmoles of G1-191-F were attached to Corning™ Costar™ Brand 96-Well EIA/RIA Plates (Fisher Scientific) and were incubated (in triplicate, overnight at 4 °C, and then at 37 °C for 1 hour with 300 rpm agitation) with eight increasing concentrations (0, 0.032, 0.16, 0.8, 4, 20, 100 and 500 nM) of each individual RNA aptamer. In turn, 10 pmoles of each 6xHis-tagged core variant (G1-191, G1-122, G2-122, G3-122, or G4-122) were attached to the corresponding wells of a Pierce® Copper Coated High Binding Capacity Plate (Thermo Scientific Fisher) and were incubated in a similar way. In both cases, four different aptamers could be assayed in parallel in a single plate. Two independent multiwell plates were used for each experiment, thus giving rise to 6 aptamer-target data points for each aptamer concentration. After the incubations, the supernatants were discarded, 100 μl/well of blocking buffer (PBS 1X with 1% BSA) was added and incubated at 37 °C for 30 min with agitation (300 rpm), and the supernatant was discarded again. As a negative control, the 76 nt-long RNA molecule termed R-ACUG, which contains 10 repeats of the tetranucleotide ACUG in its central region [5'-GCGGAUCCA GACUGGUGU(ACUG)$_{10}$GCCCUAAAGACAAGCU UC-3'] was used.

RNA aptamers bound to the corresponding variants of HCV core protein were denatured and eluted with water at 95 °C. Recovered aptamers were then amplified by RTqPCR in triplicate. The kit KAPA SYBR FAST qRT-PCR Master Mix (2X) Universal (Kapa Biosystems, Boston, Massachusetts, USA) was used for the RTqPCR amplifications. The reactions contained SYBR Green RX Mix 1X buffer, primers M1-F (0.06 μM) and M1-R (0.06 μM), KAPA RT Mix 1X and enzyme mix (0.1 U/μl). The final volume of all the reactions was 10 μl, containing 4.7 μl of the RNA template. The RTqPCR reactions were carried out in triplicate using the following cycling program: 3 min 50 °C, 5 min 95 °C, (1 min 95 °C, 1 min 55 °C, 30 sec 60 °C). The Ct (Cycle threshold) values were transformed into pmoles of bound RNA aptamer by interpolation using a calibration curve as previously described,[49] which was constructed by amplifying 10 different known concentrations (from 2.5×10^{-14} to 2.5×10^{-21} moles) of a 76 nt-long RNA molecule, in triplicate. As described for DNA aptamers, the specificity of the highly represented RNA aptamers was tested by means of ELONA-RTqPCR, using the 6xHis-tagged proteins HSA and HCV NS3.

ELONA sandwich

Biotin-labelled capture aptamer AptD-1312 (at 100 nM concentration, in SB) was added to streptavidin-coated 96-well plates (100 μl/well) and incubated at 4 °C overnight (Figure 3(B)). Then, the wells were washed three times with 200 μl of SB. G1-122 core protein was added at eight concentrations (0, 1, 2, 4, 8, 16, 32 and 64 nM, in SB) to the corresponding wells (100 μl/well, in triplicate) and incubated at 25 °C for 1 h. The wells were washed three times with 200 μl of SB and the digoxigenin-labelled detection aptamers AptD-421, AptD-1312, AptD-1932 and AptD-2251 (at 42 nM concentration, in SB) were added (100 μl/well) and incubated at 25 °C for 1 h. After washing three times with 200 μl of SB, the wells were incubated with an anti-digoxigenin antibody (Sigma) labelled with HRP (diluted 1:5000 in PBS 1X) at 25 °C for 1 h. The wells were washed three times using 200 μl of PBS 1X, and TMB (100 μl/well) was added and incubated at 25 °C for 5 to 10 min. The reactions were stopped by adding 100 μl/well of 2 M H_2SO_4 and the absorbance at 450 nm was measured in each

well using a microplate reader Tecan GENios (Tecan Group Ltd., Zürich, Switzerland). The Limit of Detection (LoD) of the assay was calculated (for n = 3) as follows: LoD = 3Sb/m, where Sb is the blank standard deviation and m is the slope of the linear regression equation.

Competitive ELONA

The competitive assay consisted of two steps (Figure 3(C)). In the first one, the digoxigenin-labelled aptamer AptD-1312 (at 5 nM concentration) was incubated with different concentrations of G1-122 core protein (0, 0.016, 0.08, 0.4, 2, 10, 50 and 250 nM in SB) in distinct Eppendorf tubes, at 37 °C for 1 h. In the second step, each mixture (composed of aptamer-core complexes, unbound aptamer and/or unbound protein) was added (100 μl/well, in triplicate) to Copper-coated 96-well plates onto which G1-122 core protein had been previously immobilized (at 100 nM) as described above. This allowed the unbound aptamer present in some of the tubes to react with the immobilized protein. After washing three times with 200 μl of SB, the wells were incubated with an anti-digoxigenin antibody labelled with HRP and the colorimetric reaction of TMB was revealed as described above. LoD was calculated as explained above.

Analysis of binding site competition in solution

A competitive, indirect ELONA was carried out for the analysis of the possible competition of aptamers in solution for the same (or overlapping) binding site in HCV core protein. For that, core G1-122 (at 100 nM concentration, in PBS 1X) was immobilized onto Copper-coated 96-well plates at 25 °C for 1 h, in triplicate. Then, the wells were washed three times using 200 μl of PBS 1X. In each experiment, a non-labelled or "cold" aptamer (AptD-421, AptD-2251 or AptD-1932) was mixed (at concentrations 0, 0.032, 0.16, 0.8, 4, 20, 100 and 500 nM) with the digoxigenin-labelled AptD-1312 at 4 nM concentration. Each mixture of aptamers (100 μl/well, in SB) was incubated at 37 °C for 1 h with the immobilized G1-122 core protein. Then, the wells were washed three times by adding 200 μl of SB, and the incubation with an anti-digoxigenin antibody labelled with HRP and the colorimetric reaction of TMB were performed as described above.

Functional analysis of high affinity aptamers in cell culture

Cells and virus origin. For the experiments in cell culture, Huh-7.5, Huh-7-Lunet and Huh-7.5 reporter cell lines were used. Their origin and the procedures for cell growth in Dulbecco's modification of Eagle's medium (DMEM), have been previously described.[83–84] Huh-7.5 cells cultured at 37 °C and under 5% CO_2 were used for aptamers transfection, infections and titration of virus infectivity. Cells were periodically thawed from a large frozen stock and passaged a maximum of 30 times.

The initial HCV strain (belonging to viral genotype 2, subtype 2a) was obtained by transcription of plasmid Jc1FLAG2(p7-nsGluc2A),[85] followed by RNA electroporation into Huh-Lunet cells, concentration of the progeny virus shed into the culture medium, and further amplification in Huh-7.5 reporter cell monolayers to yield HCV p0, as previously described.[51] The absence of contamination was controlled by titration of the supernatants of mock-infected cells, which were maintained in parallel with the infected cultures.

Transfection of aptamers and infections with HCV p0. For transfection of aptamers, 1×10^5 Huh-7.5 cells were seeded in 24-well plates. Cells were transfected 24 h later with 1 μg of aptamer AptD-1312 or AptD-1932 in SB1T buffer (see above) lacking Tween-20, which were previously refolded by incubation at 100 °C for 5 min and at 37 °C for 5 min. FuGeneHD Transfection Reagent (Promega, Madison, Wisconsin, USA) was used. The use of fluorescent microscopy and fluorescein-labelled aptamers allowed their intracellular localization within the Huh-7.5 cells, thus checking that transfection was correctly performed (Supplementary Figure S7). Twenty-four hours after transfection, cells were washed with EDTA, trypsinized and transferred to 6-well plates. Cells were infected with HCV p0 at MOI of 0.1 $TCID_{50}$/cell 24 h later and, after a virus adsorption period of 5 h at 37 °C, the inoculum was removed and 2 ml of medium was added to the cell monolayer. The infected cells were further incubated at 37 °C for 72 h.

Virus titration, RNA extraction and quantification. Titrations of infectious HCV in the presence or absence of aptamers were performed by serial dilution of cell culture supernatants that were subsequently applied to Huh-7.5 cell monolayers in 96-well plates (6,400 cells/well, seeded 16 h earlier). Three days after infection, cells were washed with PBS, fixed with ice-cold methanol, and stained to detect NS5A using anti-NS5A monoclonal antibody 9E10, as described previously.[51,86] Virus titers were expressed as $TCID_{50}$/ml.[87]

Intracellular viral RNA was extracted from infected cells using the Qiagen RNeasy kit (Qiagen, Valencia, CA, USA), following the manufacturer's instructions. In parallel, viral RNA from 140 μl of cell culture supernatants was extracted using the Qiagen QIAamp viral RNA mini kit (Qiagen, Valencia, CA, USA). In both cases, HCV RNA quantification was performed by

RTqPCR using the Light Cycler RNA Master SYBR green I kit (Roche).[51,86] Quantification was relative to a standard curve obtained with known amounts of HCV RNA synthesized by *in vitro* transcription of plasmid GNNFLAG2(p7-nsGluc2A). The 5' UTR of HCV genome was amplified using the forward primer HCV-5UTR-F2 (5'-TGAGGAACTACTGTCTT CACGCAGAAAG-3'; the 5' nucleotide corresponds to genomic residue 47 according to JFH-1, GenBank accession number AB047639) and the reverse HCV-5UTR-R2 (5'-TGCTCATGGTG CACGGTCTACGAG-3'; the 5' nucleotide corresponds to genomic residue 347 according to JFH-1).[51] Negative controls (both without template RNA and with RNA from mock-infected cells) were run in parallel with each amplification reaction to confirm the absence of contamination with undesired templates.

CRediT authorship contribution statement. **Beatriz Torres-Vázquez:** Methodology, Formal analysis, Investigation, Visualization. **Ana María de Lucas:** Methodology, Investigation. **Carlos García-Crespo:** Formal analysis, Investigation. **Juan Antonio García-Martín:** Conceptualization, Software, Formal analysis, Data curation, Visualization. **Adrián Fragoso:** Investigation. **María Fernández-Algar:** Investigation. **Celia Perales:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft. **Esteban Domingo:** Conceptualization, Resources, Writing – original draft, Supervision, Funding acquisition. **Miguel Moreno:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Supervision. **Carlos Briones:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

DATA AVAILABILITY

Data will be made available on request.

Acknowledgements

We thank the initial contributions to this research line made by David Abia, Ana García-Sacristán, Victor M. Gonzalez, Antonio Madejón, Susanna Manrubia, Elena Martín-Palma, Juan Carlos Oliveros, Javier Samaniego, and María Eugenia Soria. The work at CAB was funded by the Spanish Ministerio de Economía y Competitividad (MINECO) and EU FEDER program grants no. RTC-2014-1986-1 and BIO2016-79618-R, the Spanish Ministerio de Ciencia e Innovación (MICINN) grant no. PID2019-104903RB-I00, and the Spanish Agencia Estatal de Investigación (AEI) Project no. MDM-2017-0737 - Unidad de Excelencia "María de Maeztu". The work at CAB also benefits from the interdisciplinary framework provided by CSIC through "LifeHUB.CSIC" initiative (PIE 202120E047-Conexiones-Life). B.T-V. is supported by the predoctoral fellowship TS17/16 from INTA and by the CSIC "Garantía Juvenil" contract CAM19_PRE_CAB_001 funded by Comunidad Autónoma de Madrid (CAM). The work at CBMSO and IIS-FJD was supported by grants SAF2017-87846-R and BFU2017-91384-EXP from Ministerio de Ciencia, Innovación y Universidades (MCIU), PI18/00210 from Instituto de Salud Carlos III (ISCIII), and S2018/BAA-4370 (PLATESA2) from CAM/FEDER. C.G.-C. is supported by the predoctoral contract PRE2018-083422 from MCIU. C.P. is supported by the Miguel Servet program of ISCIII (CPII19/00001), cofinanced by the European Regional Development Fund (ERDF). CIBERehd is funded by ISCIII. Institutional grants from the Fundación Ramón Areces and Banco Santander to the CBMSO are also acknowledged. The team at CBMSO belongs to the Global Virus Network (GVN).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2022.167501.

Received 25 October 2021;
Accepted 14 February 2022;
Available online 17 February 2022

Keywords:
antivirals;
biosensors;
ELONA;
SELEX;
sequence motifs

† Current address: Centro Andaluz de Biología del Desarrollo (CABD, CSIC - U. Pablo de Olavide - Junta de Andalucía), Sevilla, Spain.

Abbreviations:
ELONA, Enzyme-Linked OligoNucleotide Assay; HCV, Hepatitis C virus; *Kd*, Dissociation constant; LoD, Limit of Detection; POC, Point-of-care; UDS, Ultra-Deep Sequencing; SELEX, Systematic Evolution of Ligands by EXponential enrichment

References

1. Rosen, H.R., (2013). Emerging concepts in immunity to hepatitis C virus infection. *J Clin Invest* **123**, 4121–4130. https://doi.org/10.1172/JCI67714.
2. Ansaldi, F., Orsi, A., Sticchi, L., Bruzzone, B., Icardi, G., (2014). Hepatitis C virus in the new era: perspectives in epidemiology, prevention, diagnostics and predictors of response to therapy. *World J Gastroenterol* **20**, 9633–9652. https://doi.org/10.3748/wjg.v20.i29.9633.
3. Shaheen, M.A., Idrees, M., (2015). Evidence-based consensus on the diagnosis, prevention and management of hepatitis C virus disease. *World J Hepatol* **7**, 616–627. https://doi.org/10.4254/wjh.v7.i3.616.
4. Manns, M.P. et al, (2017). Hepatitis C virus infection. *Nat Rev Dis Primers* **3**, 17006. https://doi.org/10.1038/nrdp.2017.6.
5. Pietschmann, T., Brown, R.J.P., (2019). Hepatitis C Virus. *Trends Microbiol* **27**, 379–380. https://doi.org/10.1016/j.tim.2019.01.001.
6. Dustin, L.B., Bartolini, B., Capobianchi, M.R., Pistello, M., (2016). Hepatitis C virus: life cycle in cells, infection and host response, and analysis of molecular markers influencing the outcome of infection and response to therapy. *Clin Microbiol Infect* **22**, 826–832. https://doi.org/10.1016/j.cmi.2016.08.025.
7. Warkad, S.D., Nimse, S.B., Song, K.S., Kim, T., (2018). HCV Detection, Discrimination, and Genotyping Technologies. *Sensors (Basel)* **18** https://doi.org/10.3390/s18103423.
8. WHO, (2017). *Global Hepatitis Report*.
9. WHO, (2021). *Global progress report on HIV, viral hepatitis and sexually transmitted infections, 2021*.
10. Luxenburger, H., Neumann-Haefelin, C., Thimme, R., Boettler, T., (2018). HCV-Specific T Cell Responses During and After Chronic HCV Infection. *Viruses* **10** https://doi.org/10.3390/v10110645.
11. Kushima, Y., Wakita, T., Hijikata, M., (2010). A disulfide-bonded dimer of the core protein of hepatitis C virus is important for virus-like particle production. *J Virol* **84**, 9118–9127. https://doi.org/10.1128/JVI.00402-10.
12. Sharma, G., Raheja, H., Das, S., (2018). Hepatitis C virus: Enslavement of host factors. *IUBMB Life* **70**, 41–49. https://doi.org/10.1002/iub.1702.
13. Domingo, E., Perales, C., (2019). Viral quasispecies. *PLOS Genetics* **15**, https://doi.org/10.1371/journal.pgen.1008271 e1008271.
14. Petruzziello, A., Marigliano, S., Loquercio, G., Cozzolino, A., Cacciapuoti, C., (2016). Global epidemiology of hepatitis C virus infection: An up-date of the distribution and circulation of hepatitis C virus genotypes. *World J. Gastroenterol.* **22**, 7824–7840. https://doi.org/10.3748/wjg.v22.i34.7824.
15. Keikha, M. et al, (2020). HCV genotypes and their determinative role in hepatitis C treatment. *Virusdisease* **31**, 235–240. https://doi.org/10.1007/s13337-020-00592-0.
16. Borgia, S.M. et al, (2018). Identification of a novel hepatitis C virus genotype from Punjab, India: expanding classification of hepatitis C virus into 8 genotypes. *J. Infect. Dis.* **218**, 1722–1729. https://doi.org/10.1093/infdis/jiy401.
17. Cuypers, L. et al, (2015). Genetic diversity and selective pressure in hepatitis C virus genotypes 1–6: significance for direct-acting antiviral treatment and drug resistance. *Viruses* **7**, 5018–5039. https://doi.org/10.3390/v7092857.
18. Wang, S.-R. et al, (2016). A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti-hepatitis C target. *Sci. Adv.* **2**, e1501535. https://doi.org/10.1126/sciadv.1501535.
19. Gallego, I. et al, (2020). Broad and dynamic diversification of infectious hepatitis C virus in a cell culture environment. *J. Virol.* **94**, e01856–01819. https://doi.org/10.1128/JVI.01856-19.
20. Kopp, M., Murray, C.L., Jones, C.T., Rice, C.M., (2010). Genetic analysis of the carboxy-terminal region of the hepatitis C virus core protein. *J. Virol.* **84**, 1666–1673. https://doi.org/10.1128/JVI.02043-09.
21. Khaliq, S., Jahan, S., Pervaiz, A., (2011). Sequence variability of HCV Core region: important predictors of HCV induced pathogenesis and viral production. *Infect. Genet. Evol.* **11**, 543–556. https://doi.org/10.1016/j.meegid.2011.01.017.
22. Mahmoudvand, S., Shokri, S., Taherkhani, R., Farshadpour, F., (2019). Hepatitis C virus core protein modulates several signaling pathways involved in hepatocellular carcinoma. *World J. Gastroenterol.* **25**, 42–58. https://doi.org/10.3748/wjg.v25.i1.42.
23. Gawlik, K., Gallay, P.A., (2014). HCV core protein and virus assembly: what we know without structures. *Immunol. Res.* **60**, 1–10. https://doi.org/10.1007/s12026-014-8494-3.
24. Strosberg, A.D., Kota, S., Takahashi, V., Snyder, J.K., Mousseau, G., (2010). Core as a novel viral target for hepatitis C drugs. *Viruses* **2**, 1734–1751. https://doi.org/10.3390/v2081734.
25. Holmstrom, E.D., Nettels, D., Schuler, B., (2018). Conformational plasticity of hepatitis C virus core protein enables RNA-induced formation of nucleocapsid-like particles. *J. Mol. Biol.* **430**, 2453–2467. https://doi.org/10.1016/j.jmb.2017.10.010.
26. Galli, C., Julicher, P., Plebani, M., (2018). HCV core antigen comes of age: a new opportunity for the diagnosis of hepatitis C virus infection. *Clin. Chem. Lab. Med.* **56**, 880–888. https://doi.org/10.1515/cclm-2017-0754.
27. Warkad, S.D., Song, K.-S., Pal, D., Nimse, S.B., (2019). Developments in the HCV screening technologies based on the detection of antigens and antibodies. *Sensors* **19** https://doi.org/10.3390/s19194257.
28. Ivanova Reipold, E. et al, (2017). Optimising diagnosis of viraemic hepatitis C infection: the development of a target product profile. *BMC Infect. Dis.* **17**, 707. https://doi.org/10.1186/s12879-017-2770-5.
29. Chevaliez, S., (2019). Strategies for the improvement of HCV testing and diagnosis. *Expert Rev. Anti Infect. Ther.* **17**, 341–347. https://doi.org/10.1080/14787210.2019.1604221.
30. NIH, (2020). *Hepatitis C*.
31. Sansonno, D. et al, (2003). Non-enveloped HCV core protein as constitutive antigen of cold-precipitable immune complexes in type II mixed cryoglobulinaemia. *Clin. Exp. Immunol.* **133**, 275–282. https://doi.org/10.1046/j.1365-2249.2003.02204.x.
32. Soria, M.E. et al, (2020). Amino acid substitutions associated with treatment failure for hepatitis C virus infection. *J. Clin. Microbiol.* **58**, 16. https://doi.org/10.1128/jcm.01985-20.
33. Di Maio, V.C. et al, (2017). Multiclass HCV resistance to direct-acting antiviral failure in real-life patients advocates

for tailored second-line therapies. *Liver Int.* **37**, 514–528. https://doi.org/10.1111/liv.13327.
34. Parczewski, M. et al, (2021). HCV resistance-associated substitutions following direct-acting antiviral therapy failure - Real-life data from Poland. *Infection, Genetics and Evolution* **93**, https://doi.org/10.1016/j.meegid.2021.104949 104949.
35. Ellington, A.D., Szostak, J.W., (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822. https://doi.org/10.1038/346818a0.
36. Tuerk, C., Gold, L., (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.
37. Bayat, P. et al, (2018). SELEX methods on the road to protein targeting with nucleic acid aptamers. *Biochimie* **154**, 132–155. https://doi.org/10.1016/j.biochi.2018.09.001.
38. Zhuo, Z. et al, (2017). Recent advances in SELEX technology and aptamer applications in biomedicine. *Int. J. Mol. Sci.* **18** https://doi.org/10.3390/ijms18102142.
39. Kaur, H., Bruno, J.G., Kumar, A., Sharma, T.K., (2018). Aptamers in the therapeutics and diagnostics pipelines. *Theranostics* **8**, 4016–4032. https://doi.org/10.7150/thno.25958.
40. Famulok, M., Mayer, G., (2014). Aptamers and SELEX in chemistry & biology. *Chem. Biol.* **21**, 1055–1058. https://doi.org/10.1016/j.chembiol.2014.08.003.
41. Liu, L.S., Wang, F., Ge, Y.H., Lo, P.K., (2021). Recent developments in aptasensors for diagnostic applications. *ACS Appl. Mater. Interfaces* **13**, 9329–9358. https://doi.org/10.1021/acsami.0c14788.
42. Zou, X.R., Wu, J., Gu, J.Q., Shen, L., Mao, L.X., (2019). Application of aptamers in virus detection and antiviral therapy. *Front. Microbiol.* **10**, 20. https://doi.org/10.3389/fmicb.2019.01462.
43. Gonzalez, V.M., Martin, M.E., Fernandez, G., Garcia-Sacristan, A., (2016). Use of aptamers as diagnostics tools and antiviral agents for human viruses. *Pharmaceuticals* **9**, 34. https://doi.org/10.3390/ph9040078.
44. Binning, J.M., Leung, D.W., Amarasinghe, G.K., (2012). Aptamers in virology: recent advances and challenges. *Front. Microbiol.* **3**, 6. https://doi.org/10.3389/fmicb.2012.00029.
45. Kim, T.H., Lee, S.W., (2021). Aptamers for anti-viral therapeutics and diagnostics. *Int. J. Mol. Sci.* **22**, 41. https://doi.org/10.3390/ijms22084168.
46. Lee, S. et al, (2007). Chip-based detection of hepatitis C virus using RNA aptamers that specifically bind to HCV core antigen. *Biochem. Biophys. Res. Commun.* **358**, 47–52. https://doi.org/10.1016/j.bbrc.2007.04.057.
47. Shi, S. et al, (2014). Inhibition of hepatitis C virus production by aptamers against the core protein. *J. Virol.* **88**, 1990–1999. https://doi.org/10.1128/jvi.03312-13.
48. Stewart, H. et al, (2016). Identification of novel RNA secondary structures within the hepatitis C virus genome reveals a cooperative involvement in genome packaging. *Sci. Rep.* **6**, 22952. https://doi.org/10.1038/srep22952.
49. Moreno, M. et al, (2019). A combined ELONA-(RT)qPCR approach for characterizing DNA and RNA aptamers selected against PCBP-2. *Molecules* **24** https://doi.org/10.3390/molecules24071213.
50. Alinovskaya, L.I., Sedykh, S.E., Ivanisenko, N.V., Soboleva, S.E., Nevinsky, G.A., (2018). How human serum albumin recognizes DNA and RNA. *Biol. Chem.* **399**, 347–360. https://doi.org/10.1515/hsz-2017-0243.
51. Perales, C. et al, (2013). Response of hepatitis C virus to long-term passage in the presence of alpha interferon: multiple mutations and a common phenotype. *J. Virol.* **87**, 7593–7607. https://doi.org/10.1128/jvi.02824-12.
52. Kikuchi, K. et al, (2005). A hepatitis C virus (HCV) internal ribosome entry site (IRES) domain III-IV-targeted aptamer inhibits translation by binding to an apical loop of domain IIId. *Nucleic Acids Res.* **33**, 683–692. https://doi.org/10.1093/nar/gki215.
53. Fukuda, K. et al, (2008). Isolation of RNA aptamers specific for the 3' X tail of HCV. *Nucleic Acids Symp. Ser.* **52**, 205–206. https://doi.org/10.1093/nass/nrn104.
54. Marton, S., Romero-López, C., Berzal-Herranz, A., (2013). RNA aptamer-mediated interference of HCV replication by targeting the CRE-5BSL3.2 domain. *J. Viral Hepatitis* **20**, 103–112. https://doi.org/10.1111/j.1365-2893.2012.01629.x.
55. Fernandez-Sanles, A. et al, (2015). RNA aptamers as molecular tools to study the functionality of the hepatitis C virus CRE region. *Molecules* **20**, 16030–16047. https://doi.org/10.3390/molecules200916030.
56. Gao, Y.M. et al, (2014). Inhibition of hepatitis C virus infection by DNA aptamer against NS2 protein. *PLoS ONE* **9**, 10. https://doi.org/10.1371/journal.pone.0090333.
57. Urvil, P.T. et al, (1997). Selection of RNA aptamers that bind specifically to the NS3 protease of hepatitis C virus. *Eur. J. Biochem.* **248**, 130–138.
58. Fukuda, K. et al, (2000). Isolation and characterization of RNA aptamers specific for the hepatitis C virus nonstructural protein 3 protease. *Eur. J. Biochem.* **267**, 3685–3694.
59. Umehara, T. et al, (2005). Rational design of dual-functional aptamers that inhibit the protease and helicase activities of HCVNS3. *J. Biochem.* **137**, 339–347. https://doi.org/10.1093/jb/mvi042.
60. Yu, X. et al, (2014). Inhibition of hepatitis C virus infection by NS5A-specific aptamer. *Antiviral Res.* **106**, 116–124. https://doi.org/10.1016/j.antiviral.2014.03.020.
61. Lee, C.H. et al, (2013). Inhibition of Hepatitis C Virus (HCV) Replication by Specific RNA Aptamers against HCV NS5B RNA Replicase. *J. Virol.* **87**, 7064. https://doi.org/10.1128/JVI.00405-13.
62. Biroccio, A., Hamm, J., Incitti, I., De Francesco, R., Tomei, L., (2002). Selection of RNA aptamers that are specific and high-affinity ligands of the hepatitis C virus RNA-dependent RNA polymerase. *J. Virol.* **76**, 3688–3696. https://doi.org/10.1128/jvi.76.8.3688-3696.2002.
63. Kanamori, H., Yuhashi, K., Uchiyama, Y., Kodama, T., Ohnishi, S., (2009). In vitro selection of RNA aptamers that bind the RNA-dependent RNA polymerase of hepatitis C virus: A possible role of GC-rich RNA motifs in NS5B binding. *Virology* **388**, 91–102. https://doi.org/10.1016/j.virol.2009.02.032.
64. Park, J.H., Jee, M.H., Kwon, O.S., Keum, S.J., Jang, S.K., (2013). Infectivity of hepatitis C virus correlates with the amount of envelope protein E2: Development of a new aptamer-based assay system suitable for measuring the infectious titer of HCV. *Virology* **439**, 13–22. https://doi.org/10.1016/j.virol.2013.01.014.
65. Yang, D. et al, (2013). Inhibition of hepatitis C virus infection by DNA aptamer against envelope protein.

Antimicrob. Agents Chemother. **57**, 4937–4944. https://doi.org/10.1128/aac.00897-13.
66. Chen, F., Hu, Y.L., Li, D.Q., Chen, H.D., Zhang, X.L., (2009). CS-SELEX generates high-affinity ssDNA aptamers as molecular probes for hepatitis C virus envelope glycoprotein E2. *PLoS ONE* **4**, 11. https://doi.org/10.1371/journal.pone.0008142.
67. Wang, C., Zhang, L., Shen, X., (2013). Development of a nucleic acid lateral flow strip for detection of hepatitis C virus (HCV) core antigen. *Nucleosides Nucleotides Nucleic Acids* **32**, 59–68. https://doi.org/10.1080/15257770.2013.763976.
68. Pleshakova, T.O. et al, (2018). The detection of hepatitis c virus core antigen using afm chips with immobolized aptamers. *J. Virol. Methods* **251**, 99–105. https://doi.org/10.1016/j.jviromet.2017.10.015.
69. Ghanbari, K., Roushani, M., (2018). A nanohybrid probe based on double recognition of an aptamer MIP grafted onto a MWCNTs-Chit nanocomposite for sensing hepatitis C virus core antigen. *Sens. Actuators, B* **258**, 1066–1071. https://doi.org/10.1016/j.snb.2017.11.145.
70. Stumper, R. et al, (2005). Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *J. Virol.* **79**, 2689–2699. https://doi.org/10.1128/jvi.79.5.2689-2699.2005.
71. Gómez-Sebastián, S., López-Vidal, J., Escribano, J.M., (2014). Significant productivity improvement of the baculovirus expression vector system by engineering a novel expression cassette. *PLoS ONE* **9**, https://doi.org/10.1371/journal.pone.0096562 e96562.
72. Altschul, S.F. et al, (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.
73. Katoh, K., Standley, D.M., (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780. https://doi.org/10.1093/molbev/mst010.
74. Rice, P., Longden, I., Bleasby, A., (2000). EMBOSS: The European molecular biology open software suite. *Trends Genet.* **16**, 276–277. https://doi.org/10.1016/s0168-9525(00)02024-2.
75. Gold, L. et al, (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS ONE* **5**, https://doi.org/10.1371/journal.pone.0015004 e15004.
76. Sanchez-Luque, F.J., Stich, M., Manrubia, S., Briones, C., Berzal-Herranz, A., (2014). Efficient HIV-1 inhibition by a 16 nt-long RNA aptamer designed by combining in vitro selection and in silico optimisation strategies. *Sci. Rep.* **4**, 10. https://doi.org/10.1038/srep06242.
77. Hoinka, J., Backofen, R., Przytycka, T.M., (2018). AptaSUITE: A Full-Featured Bioinformatics Framework for the Comprehensive Analysis of Aptamers from HT-SELEX Experiments. *Mol Ther Nucleic Acids* **11**, 515–517. https://doi.org/10.1016/j.omtn.2018.04.006.
78. Hoinka, J., Berezhnoy, A., Sauna, Z.E., Gilboa, E., Przytycka, T.M., (2014). AptaCluster - A method to cluster HT-SELEX aptamer pools and lessons from its application. *Lect. Notes Bioinforma.* **8394**, 115–128.
79. Heinz, S. et al, (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004.
80. Kuiken, C., Yusim, K., Boykin, L., Richardson, R., (2005). The Los Alamos hepatitis C sequence database. *Bioinformatics* **21**, 379–384. https://doi.org/10.1093/bioinformatics/bth485.
81. Zuker, M., (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415. https://doi.org/10.1093/nar/gkg595.
82. Kikin, O., D'Antonio, L., Bagga, P.S., (2006). QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34**, W676–W682. https://doi.org/10.1093/nar/gkl253.
83. Blight, K.J., McKeating, J.A., Rice, C.M., (2002). Highly permissive cell lines for subgenomic and genomic hepatitis C virus RNA replication. *J. Virol.* **76**, 13001–13014. https://doi.org/10.1128/jvi.76.24.13001-13014.2002.
84. Jones, C.T. et al, (2010). Real-time imaging of hepatitis C virus infection using a fluorescent cell-based reporter system. *Nat. Biotechnol.* **28**, 167–U116. https://doi.org/10.1038/nbt.1604.
85. Marukian, S. et al, (2008). Cell culture-produced hepatitis C virus does not infect peripheral blood mononuclear cells. *Hepatology* **48**, 1843–1850. https://doi.org/10.1002/hep.22550.
86. Lindenbach, B.D. et al, (2005). Complete replication of hepatitis C virus in cell culture. *Science* **309**, 623–626. https://doi.org/10.1126/science.1114016.
87. Reed, L.J., (1938). A simple method of estimating fifty per cent endpoints. *Am. J. Hygiene* **27**, 493–497.