

## STRUCTURAL DETERMINANTS THAT TARGET THE HEPATITIS C VIRUS CORE PROTEIN TO LIPID DROPLETS\*

Steeve Boulant<sup>1,2</sup>, Roland Montserret<sup>2</sup>, R. Graham Hope<sup>1</sup>, Maxime Ratinier<sup>2</sup>, Paul Targett-Adams<sup>1</sup>, Jean-Pierre Lavergne<sup>2</sup>, Francois Penin<sup>2</sup> and John McLauchlan<sup>1</sup>

From the <sup>1</sup> MRC Virology Unit, Institute of Virology, Glasgow G11 5JR, UK and  
<sup>2</sup> Institut de Biologie et Chimie des Protéines, CNRS-UMR 5086, IFR128 BioSciences,  
University of Lyon, Lyon-Gerland, Lyon F-69367, Cedex 07, France.

Running Title: Lipid Droplet Association of HCV Core Protein

Address correspondence to: John McLauchlan, MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, UK, Tel. +44 (0) 141 330 4028; Fax. +44 (0) 141 330 3520; E-Mail: [j.mclauchlan@vir.gla.ac.uk](mailto:j.mclauchlan@vir.gla.ac.uk) and Francois Penin, IBCP, CNRS/UMR5086, IFR 128 BioSciences, 7, passage du vercors, 69367 Lyon Cedex 07, France, Tel. +33 (0) 4 72 72 26; Fax. + 33 (0) 4 72 72 26; E-Mail: [f.penin@ibcp.fr](mailto:f.penin@ibcp.fr)

Hepatitis C virus core protein is targeted to lipid droplets, which serve as intracellular storage organelles, by its C-terminal domain, termed D2. From circular dichroism and nuclear magnetic resonance analyses, we demonstrate that the major structural elements within D2 consist of two amphipathic  $\alpha$ -helices (Helix I and Helix II) separated by a hydrophobic loop (HL). Both helices require a hydrophobic environment for folding, indicating that lipid interactions contribute to their structural integrity. Mutational studies revealed that a combination of Helix I, HL and Helix II is essential for efficient lipid droplet association and pointed to an in-plane membrane interaction of the two helices at the phospholipid layer interface. Aside from lipid droplet association, membrane interaction of D2 is necessary for folding and stability of core following maturation at the endoplasmic reticulum membrane by signal peptide peptidase. These studies identify critical determinants within a targeting domain that enable trafficking and attachment of a viral protein to lipid droplets. They also serve as a unique model for elucidating the specificity of protein-lipid interactions between two membrane-bound organelles.

Lipid droplets (LDs) are intracellular storage organelles, composed primarily of a hydrophobic core of cholesterol ester and triacylglycerol, which is bounded by a phospholipid leaflet encased in a layer of proteins (1). The properties of LDs have attracted considerable interest because of the link between enhanced fat storage and human diseases such as obesity, atheroma, non-alcoholic steatohepatitis and cancer. The periphery of LDs has been described as a novel membrane domain (2,3) with between 17 and 35 cellular proteins attached to the LD surface (4-6). The most widely studied LD proteins are perilipin, adipocyte differentiation-related protein (ADRP) and the tail-interacting protein, TIP47, collectively known as the PAT protein family (7). Identification of motifs within PAT proteins that are sufficient for LD association has been elusive, due to redundancy in the sequences necessary for efficient targeting (8-10). A similar picture has emerged for caveolin, which also can be found on LDs (2,11). Consequently, little is known about the processes and determinants that mediate the attachment of proteins to LDs.

Hepatitis C virus (HCV), a member of the *Flaviviridae*, is a human pathogen of global

clinical importance that typically establishes a chronic infection in infected individuals. The viral genome consists of a single-stranded, positive-sense RNA molecule, which encodes a polyprotein of some 3000 amino acids that is cleaved by cellular and viral proteases to generate the mature viral products (reviewed in 12). Structural components of the virion, which comprise the capsid protein (referred to as core) and two envelope glycoproteins (E1 and E2), lie at the polyprotein N-terminus and are released by cellular signalases at the endoplasmic reticulum (ER) membrane. E1 and E2 are produced following cleavage by signal peptidase (SP) while maturation of core requires proteolysis by SP and signal peptide peptidase (SPP; 13,14). Mature core is a dimeric,  $\alpha$ -helical protein exhibiting features that are consistent with those of a membrane protein (15). Core can be separated into two domains, D1 and D2, based on its hydrophobic profile (15,16). D1 consists of the N-terminal 117 amino acids and has similar characteristics to the capsid proteins of related pesti- and flaviviruses. The first amino acid in D2 has been positioned at residue 118 but its C-terminal limit is less well-defined and lies in the region between amino acids 171-182, which encompasses the beginning of the signal peptide between core and E1 (SP<sub>core-E1</sub>) and the SPP cleavage site. D2 is required for proper folding of domain D1 and is critical for the membrane protein characteristics of core (15). The D2 region is absent in the pesti- and flaviviruses but is found in GB virus B (GBV-B; 15,17), the closest relative to HCV in terms of sequence similarity.

Following maturation, core is targeted to LDs in a wide variety of cell types (18-20). Apart from LDs, the protein can be retained at the ER membrane (21,22) and localized on mitochondria (18,23). The purpose of the interaction between core and LDs in the virus life cycle has not been elucidated but could play a role in virus assembly or the development of liver disease in chronic infection. Our previous studies have mapped the

sequences required for association of core with LDs to segments throughout D2 (17,20). However, the lack of information concerning the structural motifs within this domain precluded any description of the nature of attachment between core and LDs.

Here, we describe structural analyses of D2, which revealed two amphipathic  $\alpha$ -helices connected by a hydrophobic loop. Mutational studies identified the essential role of several hydrophobic residues in the association of core with LDs and point to an in-plane interaction between the two helices in D2 and the surface of membranes. Moreover, we have examined the contribution of individual residues to the stability of core after SPP cleavage to assess whether there is a link between LD association and the abundance of the protein after processing. From our findings, we propose a model depicting the role of D2 structural elements in the folding of HCV core and LD targeting.

## EXPERIMENTAL PROCEDURES

*Construction of Plasmids* - Plasmids pgHCV/CE1E2<sub>Gla</sub> and pSFV/CE1E2<sub>Gla</sub> have been described previously (20,24). To express portions of HCV core protein fused to EGFP, the relevant segments of its coding region were amplified by PCR from pgHCV/CE1E2<sub>Gla</sub>. For cloning purposes, BglII and KpnI sites were introduced at the 5'- and 3'-terminal ends respectively of the primers used for amplification. Amplified DNA fragments were inserted into plasmid pEGFP-C1 (Clontech, UK) to create plasmids pGFP/D2 and the variants that consisted of sub-fragments of D2 shown in Fig. 5B. The various point mutants in Fig. 5A were generated with pgHCV/CE1E2<sub>Gla</sub> as the parent plasmid using the Quick Change Mutant kit (Stratagene). Mutagenesis oligonucleotides were designed according to manufacturer's guidelines. DNA fragments containing mutants were amplified by PCR from the relevant pgHCV/CE1E2<sub>Gla</sub> mutant plasmid and cloned into

pEGFP-C1 as described for the wild-type (wt) form of core. To generate the pSFV/CE1E2<sub>Gla</sub> mutant plasmids, a 132 bp fragment from the pgHCV/CE1E2<sub>Gla</sub> mutant plasmids, produced by cleavage by *ClaI* and *BstEII*, was ligated with appropriate *BstEII-SpeI* and *SpeI-ClaI* fragments from pSFV/CE1E2<sub>Gla</sub>. To create a construct that contained mutations at Phe<sub>130</sub> and at the SPP recognition site, a *ClaI-BstEII* fragment from pSFV/C<sub>F130E</sub>E1E2 was ligated with *BstEII-SpeI* and *SpeI-ClaI* fragments from pSFV/C<sub>spt</sub>E1E2<sub>Gla</sub> (14). For all constructs, sequences were determined to verify that only the desired mutations were introduced into DNA fragments.

*Maintenance of Tissue Culture Cells and Treatment with MG132* - HuH-7 cells were propagated in Dulbecco's modified Eagle's medium (DME) and treated with MG132 as described previously (14).

*Antibodies* - Antibodies used to detect HCV core (rabbit antisera R308), HCV E2 (monoclonal antibody [MAb] ALP98) and human ADRP have been described previously (10,20,25). A rabbit antisera against calnexin was kindly provided by Bruno Martoglio (ETH, Zurich) and anti-GFP antibody was purchased from BD Biosciences.

*Transfection of Plasmid DNA, In Vitro Transcription and Electroporation of Recombinant SFV RNA* - Plasmid DNA was transfected into HuH-7 cells using TransFast transfection reagent (Promega, UK) according to manufacturer's instructions. Cells were incubated overnight at 37°C and analyzed approximately 16 to 18 h after transfection. *In vitro* transcription and electroporation of recombinant Semliki Forest virus (SFV) RNAs into cells was performed as described previously (14,20).

*Preparation of Cell Extracts, Polyacrylamide Gel Electrophoresis and Western Blot Analysis* - To prepare extracts, cell monolayers were washed with PBS and solubilized in sample buffer (160 mM Tris-HCl, pH 6.7, 2% SDS, 700 mM  $\beta$ -mercaptoethanol, 10% glycerol and 0.004% bromophenol blue) at a concentration of

approximately  $2 \times 10^6$  cells/ml sample buffer. Samples were heated at 100°C for 5 min prior to electrophoresis through a 15% polyacrylamide gel (acrylamide:bis-acrylamide ratio of 37.5:1). For Western blot analysis, proteins, separated on polyacrylamide gels, were transferred to nitrocellulose membrane. After blocking with PBSA containing 5% milk powder (Marvel), membranes were incubated with the relevant antibody, diluted in PBSA containing 5% milk powder and 0.05% Tween-20. After washing, bound antibody was detected using protein A-peroxidase (Sigma, UK) followed by enhanced chemiluminescence.

*Separation of Intracellular Organelles on Sucrose Gradients* - HuH-7 cells, in 100 mm tissue culture dishes, were transfected with the relevant plasmid and incubated overnight at 37°C. All fractionation procedures were performed at 4°C. Cell monolayers were washed 3-4 times with PBS and once with homogenization buffer (HB; 20 mM HEPES-KOH, pH 7.4, 1 mM EDTA, 0.25 M sucrose, 1 mM PMSF). Cells were scraped into 600  $\mu$ l of HB and homogenized by passage through a 22-gauge needle. The homogenate was centrifuged at 500 g for 5 min, and the supernatant was collected as a post-nuclear supernatant (PNS). For flotation gradient centrifugation, the PNS was adjusted to 2 ml with HB containing 1.2 M sucrose and applied above a 1.6 M sucrose cushion (2 ml) in a 14x89 mm centrifuge tube. The PNS was overlaid with 2 ml each of 1, 0.75, 0.5 and 0.25 M sucrose solutions prepared in HB. Gradients were centrifuged in a SW41Ti rotor at 100,000 g for 16 h at 4°C. After centrifugation, fractions were collected, protein was precipitated from each fraction with trichloroacetic acid (15% vol/vol) and acetone (35% vol/vol) and pellets were resuspended in sample buffer.

*Indirect Immunofluorescence, Staining of Lipid Droplets and GFP Fluorescence* - Analysis of the intracellular localization of proteins and lipid droplets in live and fixed cells were performed as described in reference 10. To determine the

proportion of cells in which GFP/D2 and its variants were associated with LDs, the intracellular localization for each protein was assessed by confocal microscopy in 100 randomly selected cells.

*Peptide Synthesis and Purification* - Peptides C[118-149] (NLGKVIDTLTCGFADLMGYIPLV GAPLGGAAR) and C[141-171] (GAPLGGAAR ALAHGVRVLEDGVNYATGNLPG) were synthesized by Clonestar Biotech s.r.o. (Brno, Czech Republic) and purified by reverse phase-HPLC on a Vydac C8 column (300 Å, 10 µm, 10 X 250 mm) using a water/acetonitrile gradient containing 0.1% trifluoroacetic acid. The purity of C[118-149] and C[141-171] was greater than 98% and 90%, respectively, as determined by reverse phase HPLC, electrospray mass spectroscopy (m/z ratio of  $3204.25 \pm 0.35$  and  $2973.48 \pm 0.51$ , respectively), and nuclear magnetic resonance (NMR) spectroscopy.

*Circular Dichroism* - Circular dichroism (CD) spectra were recorded on a Jobin Yvon CD6 calibrated with 1S-(+)-10-camphorsulfonic acid. Measurements were carried out at room temperature using a 0.1 cm path length quartz cuvette (Hellma), with peptide concentrations ranging from 22 to 48 µM (pH 6.5). Spectra were recorded in the 190 to 250 nm wavelength range with a 0.2 nm increment and a 2 s integration time. Spectra were processed with CD6 software, baseline corrected, and smoothed by using a third-order least square polynomial fit. Spectral units were expressed in molar ellipticity per residue using peptide concentrations determined by measurements of UV light absorbance of tyrosine at 280 nm (molar extinction coefficient of  $1536 \text{ M}^{-1} \cdot \text{cm}^{-1}$ ). The  $\alpha$ -helical content was estimated at 222 nm using the empirical equation described in reference 26 as reported previously (27) with a theoretical molar ellipticity of  $-36000 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  for 100% helical conformation for both peptides in the various media.

*NMR Spectroscopy* - Purified peptides were dissolved in a mixture of 50% TFE-D<sub>2</sub>O (TFE-

d<sub>2</sub>, 99% isotopic enrichment) and 50% H<sub>2</sub>O (v/v) at 2.5 mM and 1.9 mM, respectively. 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added to the NMR samples as an internal <sup>1</sup>H chemical shift reference. All NMR spectra were acquired between 15°C and 25°C. Multidimensional NMR experiments were performed on a Varian Unity-plus 500 MHz equipped with a triple resonance 5 mm probe with a self-shielded z gradient coil. Nuclear Overhauser effect spectroscopy (NOESY, mixing times between 100 and 250 ms), two dimensional (2D) homonuclear total correlation spectroscopy (and clean-TOCSY, isotropic mixing time of 80 ms), and <sup>1</sup>H-<sup>13</sup>C heteronuclear single quantum correlation (HSQC) experiments were performed with conventional optimized pulse sequences as detailed previously (27-29 and references therein). Vnmr was used to process all data and Sparky (from T.D. Goddard and D.G. Kneller, University of California, San Francisco, available at URL: <http://www.cgl.ucsf.edu/home/sparky>) was employed for analysis of 2D spectra.

## RESULTS

*Domain D2 of HCV Core Protein Targets GFP to LDs* - To determine whether domain D2 could target a heterologous protein to LDs, the coding region for GFP was linked upstream from sequences encoding amino acids 118-171 from HCV core, generating a construct that produced chimeric GFP/D2 (Fig. 1). The intracellular distribution of GFP/D2 and untagged GFP in HuH-7 cells was compared with that of LDs by confocal microscopy. The locations of LDs were established by incubating cells with a BODIPY-labeled fatty acid, which accumulates in these structures following cellular uptake (10). Untagged GFP was found in both the nucleus and cytoplasm whereas GFP/D2 was detected at the periphery of LDs (Fig. 2A). We noted also that LDs had a tendency to cluster in cells expressing GFP/D2 (Fig. 2A, compare panels ii and v; data

not shown). A similar phenomenon was observed in cells expressing the full-length core protein (20). We assume that the association of core and GFP/D2 with LDs contributes to the change in their characteristics. To provide additional biochemical evidence that GFP/D2 behaved similarly to LD-associated proteins, extracts from transfected cells were fractionated by flotation centrifugation on sucrose gradients (Fig. 2B). Individual fractions were probed with antibodies against GFP, ADRP, a marker for LDs (10), and calnexin, a marker for ER membranes (30), to determine the mobility profiles of these various proteins in gradients. GFP/D2 displayed an identical mobility profile to ADRP, which differed from that for GFP alone (Fig. 2B). Compared with GFP/D2 and ADRP, calnexin was detected only in the lower portion of the gradient. These data provided further conclusive evidence that GFP/D2 associated with LDs. Thus, D2 of core contains targeting determinants that are sufficient for directing proteins to LDs.

*Structural Analysis Identifies Two  $\alpha$ -Helices within D2* – Recently, we have shown that full-length core, composed of D1 and D2, is mainly  $\alpha$ -helical (15). As a first stage to identify the locations of structural elements within D2, its sequence was analyzed by a suite of secondary structure prediction programs using euHCVdb database facilities (URL: <http://euhcvdb.ibcp.fr>; 31). The predictions suggested that two  $\alpha$ -helical segments were present in D2. One helix was predicted in region 144-161 by all of the methods whereas only some programs identified a second helix in region 118-135 (unpublished data). Other reports have predicted  $\alpha$ -helices between positions 116-134 (32) and 144-165 (11). It has been suggested also that the second helix exists as two shorter  $\alpha$ -helices between residues 145-152 and 155-159 (32). To gain experimental evidence for the existence of the predicted helices and to clarify their positions within D2, we analyzed the secondary structures of two overlapping, synthetic

peptides, C[118-149] and C[141-171], that encompassed the sequences needed for efficient targeting to LDs, by CD and NMR.

a) CD Analysis - Peptide C[118-149] was poorly soluble in water and gave a complex spectrum with a large negative band around 200 nm and a shoulder at about 220 nm (Fig. 3A), indicating a mixture of random coil structure with some poorly defined secondary structures. As an alternative to aqueous solution, TFE is known to stabilize the folding of peptidic sequences with an intrinsic propensity to adopt an  $\alpha$ -helical structure (33,34). Dissolving the peptide in 50% TFE generated a spectrum that was typical of  $\alpha$ -helical folding with two minima at 208 and 222 nm and a maximum at 192 nm (Fig. 3A and B). Titration of C[118-149] peptide folding with increasing proportions of TFE gave spectra that were characteristic for  $\alpha$ -helices above 20% TFE (unpublished data). Maximal amplitudes were reached at 45% TFE and no change was found at higher TFE proportions of up to 75% (unpublished data). The results were consistent with the suggestion that, for a peptide with helical propensity, helicity is generally evident at 20-30% TFE and folding is complete by 50% (35).

In the presence of various detergents used to mimic a membranous environment, CD spectra of C[118-149] peptide again showed the typical shape of  $\alpha$ -helix folding (Fig. 3A and B). This behavior was indicative of the capacity of the peptide to adopt an  $\alpha$ -helical structure on binding to lipid-like molecules. Interestingly, except for LPC, the amplitude maxima in CD spectra were higher in the various detergents than those observed in 50% TFE (Fig. 3A and B). Quantitatively, about 40% helical content within the peptide was estimated at 222 nm in the presence of DM and DPC; in 50% TFE, this figure dropped to about 30%. Such different values were consistent with improved stabilization of the helical region of C[118-149] peptide in a micellar context. The amplitude variations observed in the

various detergents could be attributed to altered peptide interactions with the different polar head groups of detergents. In summary, CD spectral analysis indicated the high propensity of C[118-149] to bind to lipids and adopt an  $\alpha$ -helical structure upon binding.

Peptide C[141-171] was very soluble in water and exhibited a typical CD spectrum of random structure, as denoted by the large negative band around 200 nm (Fig. 3C). Addition of TFE induced helical folding of the peptide (Fig. 3C and D), although the shape of the spectrum was less typical of an  $\alpha$ -helix fold and helical content was limited to 26% when estimated at 222 nm. In the presence of detergents, CD spectra of C[141-171] did show the characteristic shape of  $\alpha$ -helix folding with double minima at 208 and 222 nm (Fig. 3C and D). In contrast to C[118-149], amplitude maxima were lower in the various detergents than those observed in 50% TFE, except for DPC. Analysis of  $\alpha$ -helix content for the various detergents at different pH values indicated that electrostatic interactions between the peptide and the polar head of detergents were important for stability of the C[141-171]  $\alpha$ -helical fold (unpublished data). In summary, as with C[118-149], peptide C[141-171] showed a clear but more limited propensity to adopt an  $\alpha$ -helical structure upon binding to lipid-like molecules.

b) NMR Analysis - To gain detailed information at the atomic level of secondary structures formed by the peptides, their conformations were investigated by NMR. 50% TFE was considered to be an appropriate medium for such analysis since approximately the same  $\alpha$ -helical content was observed for the peptides in TFE and in detergents. 2D homo- and heteronuclear NMR analyses of the peptides in 50% TFE- $d_2$  yielded well-resolved spectra (unpublished data), which allowed sequential assignment of the spin systems by standard methodology (36). Fig. 4A gives an overview of the sequential and medium range NOE connectivities. NOE connectivity patterns

clearly indicated that the main body of both peptides displayed typical characteristics of an  $\alpha$ -helix, including strong  $dN(i,i+1)$  and medium  $d\alpha N(i, i+1)$  sequential connectivities, weak  $d\alpha N(i, i+2)$ , medium  $d\alpha N(i, i+3)$ , medium or strong  $d\alpha\beta(i, i+3)$ , and weak  $d\alpha N(i, i+4)$  medium range connectivities. For peptide C[118-149],  $\alpha$ -helical NOE connectivities were observed between amino acids L119 and I137, although some  $d\alpha N(i,i+3)$  and  $d\alpha\beta(i,i+3)$  signals were weaker than expected for a canonical helix. Such a pattern of connectivities could denote some flexibility in the helix. In the remaining C-terminal sequence, which included Pro and Gly residues, only a few medium range NOE contacts were observed, indicating that this region was mainly unstructured. For peptide C[141-171], numerous connectivities characteristic of an  $\alpha$ -helix were detected between A148 and A165. Again, some connectivities were weaker than expected, suggesting some local flexibilities in the helix. Both termini of the helix (G141-A147 and T166-G171) gave almost no medium range connectivities and therefore appeared to be essentially unstructured. Differences of  $^1H\alpha$  chemical shifts from those found in a random coil conformation are additional indicators of secondary structure (37). Relative to a random conformation, an increase of helicity results in a negative variation of the  $^1H\alpha$  chemical shifts ( $\Delta\delta^1H\alpha$ ) equal to or larger than 0.1 ppm. The long series of negative  $^1H\alpha$  chemical shift differences ( $\Delta\delta^1H\alpha \leq -0.1$  ppm) observed for segment 119-135 for peptide C[118-149] and segment 147-159 for peptide C[141-171] were typical of an  $\alpha$ -helical conformation (Fig. 4B). For peptide C[141-171], V162 to Y164 also were likely folded into the  $\alpha$ -helix, based on presence of the expected medium range NOE connectivities and the slightly positive  $\Delta\delta^1H\alpha$  observed for residues D160 and G161, which probably indicate some flexibility within the  $\alpha$ -helix. In summary, by combining the data of NOE connectivities and chemical shift

indices, we deduced that  $\alpha$ -helices extend from L119 to Y136, and from A148 to Y164. These helices were termed Helix I and II respectively.

From  $\alpha$ -helix projections, both helices were amphipathic with all charged and polar residues located on one face while the hydrophobic residues were clustered on the opposite side (Fig. 4C). The hydrophobic face of Helix I exhibited a higher number of large hydrophobic residues (Leu, Phe, Ile) compared to Helix II. The C-terminus of both helices ended with a conserved Tyr residue at the hydrophobic-hydrophilic interface, a feature that is typical of helices, which interact at the membrane interface (38). Hence, amino acids on the hydrophobic faces of Helix I and II are likely to be important for membrane binding.

Outside of Helix I and II, segment 137-147 was unstructured from the NMR data. Residues A148 and R149 in Helix II were unstructured in peptide C[118-149] but not in C[141-171]. This discrepancy was due to the absence of conformation stabilization by the hydrogen bond helix network in C[118-149] because of the location of these amino acids at the end of the peptide. Thus, Helix I and II are separated by a region with no apparent secondary structure (amino acids 137-147), which is composed of hydrophobic residues and contains two prolines. This segment could form a connecting hydrophobic loop (HL) between Helix I and II (Figs. 4B and 5A). The remainder of D2 from amino acids 165-171 appeared to be random.

*Contribution of Helix I, HL and Helix II to LD Association* - Following on from the structural analysis, we tested the ability of fragments of D2 linked to GFP to associate with LDs to try to define a minimal targeting motif. GFP fused to amino acids 118-164 of core was detected at the surface of LDs with equal efficiency to GFP/D2 (Fig. 5B and C). Hence, a combination of Helix I, HL and Helix II was sufficient for LD association. The three C-terminal amino acids (V162, N163 and Y164) that reside immediately beyond the

proposed flexible region of Helix II were dispensable for efficient LD attachment since their removal had no impact on GFP localization (Fig. 5B). Deletion of additional segments from D2 to either amino acid 149 or 141 reduced but did not abolish the association of GFP with LDs (Fig. 5B and C), indicating that Helix I along with the N-terminal part of HL retained some limited capacity for LD attachment. However, truncating D2 to position 135 gave a GFP chimeric protein that was indistinguishable from untagged GFP (Fig. 5B). Thus, Helix I alone could not direct a protein to LDs. To determine whether HL combined with Helix II retained targeting, two chimeric products, GFP/135-161 and GFP/135-171, were expressed but neither associated with LDs (Fig. 5B and C). Thus, efficient targeting to LDs was not dependent solely on a single structural motif but relied on a combination of Helix I, HL and Helix II. However, not all of these elements apparently made equal contributions since residues 118-141 and 118-149 partially targeted LDs but combinations of HL and Helix II did not direct proteins to LDs. These data indicate that Helix I and the N-terminal part of HL played a particularly critical role in LD association.

*Residues on the Hydrophobic Face of the Helices are Critical for LD Association* - To examine whether individual amino acids in the two helices and HL influenced LD targeting, a series of constructs was generated that contained point mutations in GFP/D2 between positions 119 and 158 (Fig. 5A) The positions of mutations were selected by taking into account the high degree of amino acid conservation in all genotypes (15). Since Helix I and HL appeared to play a more important role in LD targeting, a greater number of mutations were introduced into these segments, especially at the hydrophobic positions. Residues in Helix I and II were altered to glutamic acid since this amino acid is highly hydrophilic but less likely to disrupt an  $\alpha$ -helical conformation. Indeed, among the 20 amino acids, glutamic acid exhibits the highest probability to be involved in

helical structure (39). The mutations were inserted into pGFP/D2 and the intracellular distribution of the resultant GFP/D2 variants was assessed by confocal microscopy (Fig. 5C and Table 1).

The mutations decreased LD targeting to varying degrees and, from a series of experiments, two distinct patterns of mutant GFP/D2 fluorescence were observed, which were used to characterize their behavior (Table 1). Class A represented cells with evidence of GFP fluorescence at the surface of the LDs (Fig. 5C, GFP/D2<sub>C128E</sub>); any mutants with particularly weak LD targeting (Fig. 5C, GFP/D2<sub>L126E</sub>) were noted (indicated by a star in Table 1). Class B represented cells with no GFP fluorescence surrounding LDs (Fig. 5C, GFP/D2<sub>F130E</sub>). The profiles of mutants with weak or no association with LDs did differ from that for GFP/D2 by biochemical fractionation (supplemental Fig. S1). However, the inability to readily separate fractions containing LDs from those containing membranes precluded using this approach to assess subtle differences between mutants.

For amino acid changes on the hydrophobic side of Helix I, mutation of residues Leu<sub>119</sub> and Ile<sub>123</sub> did not substantially decrease LD association (Table 1). Substitution of Leu<sub>126</sub> reduced the proportion of cells with LD association for GFP/D2 to 50% and the fluorescence around LDs was weak (Fig. 5C, GFP/D2<sub>L126E</sub>). Particularly noteworthy was mutation of Phe<sub>130</sub>, which completely abolished attachment to LDs by GFP/D2 (Fig. 5C, GFP/D2<sub>F130E</sub>). Changing Leu<sub>133</sub> and Met<sub>134</sub> decreased LD distribution to 20% and 75% of cells examined respectively (Table 1). Moreover, both mutations gave weak targeting to LDs, similar to the Leu<sub>126</sub> mutant. By contrast, substitution of Lys<sub>121</sub> and Cys<sub>128</sub> on the hydrophilic side of Helix I had no effect on LD association (Fig. 5C, GFP/D2<sub>C128E</sub>; Table 1). The result with the Cys<sub>128</sub> mutant is interesting, since it lies between critical residues at positions 126 and 130. From these data, hydrophobic amino acids towards the C-terminus

of Helix I were critical for association of D2 to LDs whereas those on the hydrophilic face apparently did not participate in attachment.

In HL, mutation of Leu<sub>139</sub> partially impaired LD attachment while there was no effect on association by changing Gly<sub>145</sub> (Table 1). By contrast, mutation at three other amino acids in HL (Val<sub>140</sub>, Leu<sub>144</sub> and Ala<sub>147</sub>) reduced the proportion of cells with LD association to 30-50% and fluorescence around LDs was weak. Along with previously published data on the effect of mutations at proline residues positioned at amino acids 138 and 143 (17), our data indicate that although HL did not have any apparent secondary structure in the context of the synthetic peptides, it contained amino acids that play an essential role in efficient LD association.

Among the four amino acid mutations generated in Helix II, mutation at only Leu<sub>151</sub> on the hydrophobic face sharply reduced LD association to 40% and the fluorescence around LDs was weak (Table 1). Other mutations on the hydrophobic side of Helix II either slightly reduced (90% of cells had weak LD association for Ala<sub>150</sub>) or had no effect on LD targeting (Leu<sub>158</sub>). Similar to the results obtained with Helix I, a mutation on the hydrophilic face of Helix II did not impair LD attachment (Arg<sub>149</sub>). Therefore, in agreement with the results obtained from expressing fragments of D2 linked to GFP, Helix II did contribute to LD association but played a less critical role compared to Helix I.

To conclude, our results demonstrate that hydrophobic, but not hydrophilic, amino acids within Helix I, HL and Helix II, are critical for LD association. This suggests direct interaction of the hydrophobic residues in both helices and HL with the hydrophobic tails of phospholipids in LDs probably occurs. Due to their amphipathic nature, Helix I and II are likely to bind in-plane with the membrane in a similar fashion to that described recently for the HCV NS5A membrane anchor (40).

*Hydrophobic Amino Acids in D2 are Essential for Stability of Core Protein* - To examine whether mutations at the hydrophobic residues studied in the GFP/D2 system exerted any other influence on the properties of core, they were introduced into full-length HCV core protein and expressed in a SFV vector that produces the HCV structural components (Fig. 1).

Our results revealed that either low or undetectable levels of core were found for most mutations at hydrophobic residues (Fig. 6A). To address whether the low levels of protein resulted from degradation, experiments were performed in the presence and absence of MG132, a proteasome inhibitor (Fig. 6A). Apart from Leu<sub>119</sub>, mutation at hydrophobic amino acids gave much greater levels of core on addition of MG132 to cells. Indeed, core could be detected only in the presence of MG132 for mutations at residues Leu<sub>126</sub>, Phe<sub>130</sub>, Leu<sub>133</sub>, Met<sub>134</sub>, Leu<sub>144</sub>, Ala<sub>147</sub>, Leu<sub>151</sub> and Leu<sub>158</sub>. By contrast, mutation at hydrophilic amino acids (K<sub>121</sub>, Cys<sub>128</sub>, Gly<sub>145</sub> and Arg<sub>149</sub>) produced roughly similar amounts of core in the presence and absence of the inhibitor and therefore behaved in an equivalent manner to wt protein (Fig. 6A). Similar levels of E2 glycoprotein were made by each construct and addition of MG132 had little or no effect on its abundance (unpublished data); these findings agreed with our previous studies (14,20). This pattern of expression was observed also from indirect immunofluorescence analysis of cells electroporated with SFV RNA from all of the constructs (Fig. 6B). In MG132-treated cells, the most prominent feature for the mutants that produced unstable core was increased fluorescence in punctate regions of the nucleus but there was no evidence of LD association (Fig. 6B, panel xi). By contrast, mutations at the hydrophilic residues did not disrupt LD attachment (Fig. 6B, panels v and vii). We concluded from these data that the hydrophobic residues in Helix I, HL and Helix II were required for stability of core, therefore extending their role beyond solely a function in LD targeting.

To investigate whether degradation occurred before or after core maturation, mutations at the SPP recognition site in SP<sub>core-E1</sub> in plasmid pSFV/C<sub>sppm</sub>E1E2 (see legend of Fig. 1) were introduced into the pSFV/C<sub>F130E</sub>E1E2 construct. In a previous report, these mutations blocked SPP cleavage and enhanced the stability of core lacking amino acids 125-144 (14). Core protein, which contained the Phe<sub>130</sub> mutation but was not processed by SPP, was identified readily in the absence of MG132; by contrast, core containing this mutation, which was cleaved by SPP, was detected at much lower levels (Fig. 6C). Therefore, degradation of unstable forms of core occurs after processing by SPP.

## DISCUSSION

In this report, we have investigated the properties of D2, a domain in HCV core that is absent in the equivalent proteins of related flaviviruses, but present in GBV-B. Our studies have identified structural elements within D2, which enable efficient attachment of core to LDs, and we have demonstrated that hydrophobic amino acids within these elements are critical for targeting (summarized in Fig. 7). Moreover, hydrophobic residues in D2 perform a secondary purpose to prevent degradation of core upon its maturation (Fig. 7), probably by promoting correct folding of the protein through association with membranes. This dual purpose for D2 provides new insights into the folding and targeting of HCV core and the nature of interactions between proteins and LDs.

Our structural analysis of D2 revealed two  $\alpha$ -helices (Helix I and II), which lie between amino acids 119-136 and 148-164. Both helices are amphipathic with the hydrophilic and hydrophobic amino acids aligned on alternate faces. HL, the region between the helices, is composed primarily of hydrophobic residues and has no apparent secondary structure in the synthetic peptides that were analyzed. We

consider it likely that the hydrophobic amino acids on both helices and those in HL form a contiguous stretch of hydrophobic residues, which interact in-plane with membranous interfaces. Such an interaction is in keeping with the presence of only one layer of phospholipid at the LD surface. From the data obtained with synthetic peptides, a hydrophobic environment is crucial for promoting formation of the helices in D2, suggesting that interaction with membranes is necessary for the folding and structural integrity of the domain. In related studies, we demonstrated that the ability of full-length core to adopt an  $\alpha$ -helical conformation was dependent on both the presence of D2 and detergent (15). We deduce that membrane-binding is not only important for formation of Helix I and II in D2 but this process also plays an integral part in the folding of the entire protein. Therefore, during translation and translocation of the HCV polyprotein, interaction of D2 with the ER membrane is likely to promote folding of Helix I and II, and thereby the remainder of core (Fig. 8, stages i and ii).

Release of core from the viral polyprotein requires two cleavage events at the SP<sub>core-E1</sub> signal sequence, firstly by SP and then SPP (Fig. 8, stages i and iii). After maturation by SPP, core is directed from the ER membrane to LDs. From the evidence in our study, mutation of hydrophobic amino acids in Helix I, HL and Helix II can lead to an alternative pathway in which core is degraded upon maturation. SPP cleavage is actively engaged in degradation since blocking this process by mutation of the signal sequence prevents loss of core. The range of functions performed by SPP has not been fully determined but it is needed for generating epitopes that are presented by HLA-E to natural killer cells via the proteasome (41,42). Moreover, this protease can recognise incorrectly assembled transmembrane domains and it has been suggested to play a role in monitoring the quality of ER membrane-bound proteins (43). Based on our findings, the association with membranes mediated by residues in both helices and HL could

be pivotal in determining the fate of core. Changes in hydrophobic residues that interfere with membrane engagement could perturb the folding of D2 and thereby that of core. Under such circumstances, cleavage by SPP could prompt targeting to the proteasome for degradation (Fig. 8).

For correctly-folded and membrane-bound core, the mechanism for its transfer from ER to LDs is not known. We favour a process whereby core diffuses along the ER until it encounters a LD closely apposed or attached to the membrane (Fig. 8, stages iv and v; 10). Transfer of core between the two organelles is likely to be governed by the greater affinity of determinants in D2 for the surface of LDs as compared to the ER membrane. The minimal targeting motif for efficient LD association is a combination of Helix I, HL and Helix II. We consider that the in-plane membrane interaction of this motif is a crucial factor that enables transfer of core to the monolayer of phospholipid at the LD surface. Movement of proteins between different membrane-bound organelles can be controlled by post-translational modifications such as in the case of Rab proteins (44). We have no evidence for modifications in D2 that affect localisation but propose that specificity is determined by the characteristics of the residues and their configuration within the helices and the intervening HL sequence. Affinity also is probably determined by differences in the lipid compositions at the surface of the ER membrane and LDs. It has been reported that the phospholipid layer of LDs has a fatty acid composition that is distinct from that of the ER (3). Such differences between the two organelles could modulate the structure of D2 and its interactions with lipids to favour association with LDs rather than the ER membrane.

Aside from HCV core, LD targeting sequences have been studied for caveolin, oleosin and members of the PAT family of proteins. For the PAT proteins, sequences required to direct ADRP to LDs are distributed along the length of

the polypeptide and no contiguous targeting element has been identified (9,10). In the case of perilipin, a central region spanning about 130 amino acids, which contains three distinct hydrophobic sequences, is necessary for LD association (8,45). However, removal of any one of these three elements does not disrupt LD attachment, again demonstrating the redundancy in targeting sequences. In the case of caveolin-1 and -2, a central hydrophobic domain is required for LD targeting (11). No specific amino acids in this region in caveolin-1 could be mutated that abolished LD association and fusion of the hydrophobic domain of caveolin-2 to GFP did not direct the chimeric protein to LDs (4,11). It has been suggested that the central domain of caveolin-1 contains two hydrophobic helices and correct packing of these helices promotes LD association rather than any specific role for individual residues (11). Oleosin also contains a central hydrophobic domain, which is a key element for LD attachment. From CD and FTIR analysis, the hydrophobic domain is thought to consist of an  $\alpha$ -helical hairpin structure with a turn that is introduced by two prolines, which are separated by three hydrophobic amino acids (46-48).

As we have suggested previously, the configuration of elements in D2 of HCV core most closely matches that of oleosin compared with any of the mammalian LD proteins (17). However, there are distinct differences between D2 and the oleosin hydrophobic domain. Firstly, the targeting region in the plant protein has greater

hydrophobicity than D2. Secondly, the helical regions flanking the turn introduced by the prolines in oleosin are interchangeable without complete loss of LD targeting, which implies that alterations in the amino acid sequences of the helices may be tolerated (49). Additional structural analysis of regions involved in LD association will further our understanding of the nature of protein-lipid interactions that determine specificity for these organelles.

The lack of consistency in identifying a common set of structural motifs in LD-associated proteins may suggest that the nature of interactions between proteins and these organelles can be diverse. The diversity could reflect functional characteristics of each protein linked to their purpose in interacting with LDs. In the case of core, it is possible that the protein binds to LDs to facilitate assembly of the virus. With the advent of a system that allows propagation of HCV (50-52) and evidence that core protein associates with LDs in infected cells (53), experiments are in progress to test this hypothesis.

In conclusion, the characterization of structural determinants within the D2 domain, along with extensive mutational studies, has provided insight into the mechanisms that promote folding of HCV core and its targeting to LDs. Our study also provides a new model for examining the interaction between proteins and membrane interfaces, and the factors which govern specificity of proteins for distinct membrane-bound intracellular organelles.

## REFERENCES

1. Murphy, D.J. (2001) *Prog. Lipid Res.* **40**, 325-438
2. Fujimoto, T., Kogo, H., Ishiguro, K., Tauchi, K., and Nomura, R. (2001) *J. Cell Biol.* **152**, 1079-1085
3. Tauchi-Sato, K., Ozeki, S., Houjou, T., Taguchi, R., and Fujimoto, T. (2002) *J. Biol. Chem.* **277**, 44507-44512
4. Fujimoto, Y., Itabe, H., Sakai, J., Makita, M., Noda, J., Mori, M., Higashi, Y., Kojima, S., and Takano, T. (2004) *Biochim. Biophys. Acta* **1644**, 47-59

5. Umlauf, E., Csaszar, E., Moertelmaier, M., Schuetz, G.J., Parton, R.G., and Prohaska, R. (2004) *J. Biol. Chem.* **279**, 23699-23709
6. Brasaemle, D.L., Dolios, G., Shapiro, L., and Wang, R. (2004) *J. Biol. Chem.* **279**, 46835-46842
7. Lu, X., Gruia-Gray, J., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Londos, C., and Kimmel, A.R. (2001) *Mamm. Genome* **12**, 741-749
8. Garcia, A., Sekowski, A., Subramanian, V., and Brasaemle, D.L. (2003) *J. Biol. Chem.* **278**, 625-635
9. Nakamura, N., and Fujimoto, T. (2003) *Biochem. Biophys. Res. Commun.* **306**, 333-338
10. Targett-Adams, P., Chambers, D., Gledhill, S., Hope, R.G., Coy, J.F., Girod, A., and McLauchlan, J. (2003) *J. Biol. Chem.* **278**, 15998-16007
11. Ostermeyer, A.G., Ramcharan, L.T., Zeng, Y., Lublin, D.M., and Brown, D.A. (2004) *J. Cell Biol.* **164**, 69-78
12. Penin, F., Dubuisson, J., Rey, F.A., Moradpour, D., and Pawlotsky, J.M. (2004) *Hepatology* **39**, 5-19
13. Hussy, P., Langen, H., Mous, J., and Jacobsen, H. (1996) *Virology* **224**, 93-104
14. McLauchlan, J., Lemberg, M.K., Hope, G., and Martoglio, B. (2002) *EMBO J.* **21**, 3980-3988
15. Boulant, S., Vanbelle, C., Ebel, C., Penin, F., and Lavergne, J.P. (2005) *J. Virol.* **79**, 11353-11365
16. McLauchlan, J. (2000) *J. Viral Hepat.* **7**, 2-14
17. Hope, R.G., Murphy, D.J., and McLauchlan, J. (2002) *J. Biol. Chem.* **277**, 4261-4270
18. Moradpour, D., Englert, C., Wakita, T., and Wands, J.R. (1996) *Virology* **222**, 51-63
19. Barba, G., Harper, F., Harada, T., Kohara, M., Goulinet, S., Matsuura, Y., Eder, G., Schaff, Z., Chapman, M.J., Miyamura, T., and Brechot, C. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1200-1205
20. Hope, R.G., and McLauchlan, J. (2000) *J. Gen. Virol.* **81**, 1913-1925
21. Santolini, E., Migliaccio, G., and La Monica, N. (1994) *J. Virol.* **68**, 3631-3641
22. Lo, S.Y., Masiarz, F., Hwang, S.B., Lai, M.M., and Ou, J.H. (1995) *Virology* **213**, 455-461
23. Schwer, B., Ren, S., Pietschmann, T., Kartenbeck, J., Kaehlcke, K., Bartenschlager, R., Yen, T.S., and Ott, M. (2004) *J. Virol.* **78**, 7958-7968
24. Patel, J., Patel, A.H., and McLauchlan, J. (1999) *J. Gen. Virol.* **80**, 1681-1690
25. Clayton, R.F., Owsianka, A., Aitken, J., Graham, S., Bhella, D., and Patel, A.H. (2002) *J. Virol.* **76**, 7672-7682
26. Chen, Y.H., Yang, J.T., and Chau, K.H. (1974) *Biochemistry* **13**, 3350-3359
27. Montserret, R., McLeish, M.J., Bockmann, A., Geourjon, C., and Penin, F. (2000) *Biochemistry* **39**, 8362-8373
28. Penin, F., Geourjon, C., Montserret, R., Bockmann, A., Lesage, A., Yang, Y.S., Bonod-Bidaud, C., Cortay, J.C., Negre, D., Cozzone, A.J., and Deleage, G. (1997) *J. Mol. Biol.* **270**, 496-510
29. Favier, A., Brutscher, B., Blackledge, M., Galinier, A., Deutscher, J., Penin, F., and Marion, D. (2002) *J. Mol. Biol.* **317**, 131-144
30. Wada, I., Rindress, D., Cameron, P.H., Ou, W.J., Doherty 2<sup>nd</sup>, J.J., Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y., and Bergeron, J.J. (1991) *J. Biol. Chem.* **266**, 19599-19610
31. Combet, C., Penin, F., Geourjon, C., and Deleage, G. (2004) *Appl. Bioinformatics* **3**, 237-240

32. Suzuki, R., Sakamoto, S., Tsutsumi, T., Rikimaru, A., Tanaka, K., Shimoike, T., Moriishi, K., Iwasaki, T., Mizumoto, K., Matsuura, Y., Miyamura, T., and Suzuki, T. (2005) *J. Virol.* **79**, 1271-1281
33. Buck, M. (1998) *Q. Rev. Biophys.* **31**, 297-355
34. Montserret, R., Aubert-Foucher, E., McLeish, M.J., Hill, J.M., Ficheux, D., Jaquinod, M., van der Rest, M., Deleage, G., and Penin, F. (1999) *Biochemistry* **38**, 6479-6488
35. Jasanoff, A., and Fersht, A.R. (1994) *Biochemistry* **33**, 2129-2135
36. Wuthrich, K. (1986) John Wiley & Sons, New York
37. Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L., and Sykes, B.D. (1995) *J. Biomol. NMR* **6**, 135-140
38. Granseth, E., von Heijne, G., and Elofsson, A. (2005) *J. Mol. Biol.* **346**, 377-385
39. Chou, P.Y and Fasman G.D. (1974) *Biochemistry* **13**, 211-222
40. Penin, F., Brass, V., Appel, N., Ramboarina, S., Montserret, R., Ficheux, D., Blum, H.E., Bartenschlager, R., and Moradpour, D. (2004) *J. Biol. Chem.* **279**, 40835-40843
41. Lemberg, M.K., Bland, F.A., Weihofen, A., Braud, V.M., and Martoglio, B. (2001) *J. Immunol.* **167**, 6441-6446
42. Bland, F.A., Lemberg, M.K., McMichael, A.J., Martoglio, B., and Braud, V.M. (2003) *J. Biol. Chem.* **278**, 33747-33752
43. Crawshaw, S.G., Martoglio, B., Meacock, S.L., and High, S. (2004) *Biochem. J.* **384**, 9-17
44. Seabra, M.C., and Wasmeier, C. (2004) *Curr. Opin. Cell Biol.* **16**, 451-457
45. Subramanian, V., Garcia, A., Sekowski, A., and Brasaemle, D.L. (2004) *J. Lipid Res.* **45**, 1983-1991
46. Millichip, M., Tatham, A.S., Jackson, F., Griffiths, G., Shewry, P.R., and Stobart, A.K. (1996) *Biochem. J.* **314**, 333-337
47. Lacey, D.J., Wellner, N., Beaudoin, F., Napier, J.A., and Shewry, P.R. (1998) *Biochem. J.* **334**, 469-477
48. Alexander, L.G., Sessions, R.B., Clarke, A.R., Tatham, A.S., Shewry, P.R., and Napier, J.A. (2002) *Planta* **214**, 546-551
49. Abell, B.M., Hahn, M., Holbrook, L.A., and Moloney, M.M. (2004) *Plant J.* **37**, 461-70
50. Lindenbach, B.D., Evans, M.J., Syder, A.J., Wolk, B., Tellinghuisen, T.L., Liu, C.C., Maruyama, T., Hynes, R.O., Burton, D.R., McKeating, J.A., and Rice, C.M. (2005) *Science* **309**, 623-626
51. Wakita, T., Pietschmann, T., Kato, T., Date, T., Miyamoto, M., Zhao, Z., Murthy, K., Habermann, A., Krausslich, H.G., Mizokami, M., Bartenschlager, R., and Liang, T.J. (2005) *Nat. Med.* **11**, 791-796
52. Zhong, J., Gastaminza, P., Cheng, G., Kapadia, S., Kato, T., Burton, D.R., Wieland, S.F., Uprichard, S.L., Wakita, T., and Chisari, F.V. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9294-9299
53. Rouille, Y., Helle, F., Delgrange, D., Roingeard, P., Voisset, C., Blanchard, E., Belouzard, S., McKeating, J., Patel, A.H., Maertens, G., Wakita, T., Wychowski, C., and Dubuisson, J. (2006) *J. Virol.* **80**, 2832-2841
54. Merutka, G., Dyson, H.J., and Wright, P.E. (1995) *J. Biomol. NMR.* **5**, 14-24

## FOOTNOTES

\* This work was supported by the UK Medical Research Council (JMcL), CNRS and University of Lyon, grants from ANRS (to J-P. L and F.P.) and from the ATC “hepatites” program of INSERM (F. P.).

<sup>1</sup>The abbreviations used are: CD, circular dichroism; DM, n-dodecyl  $\beta$ -D-maltoside; DPC, dodecyl phosphocholine; HCV, hepatitis C virus; LD(s), lipid droplets(s); LPC, L- $\alpha$ -lysophosphatidylcholine; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; SFV, Semliki Forest virus; TFE, 2,2,2-trifluoroethanol; wt, wild-type.

## FIGURE LEGENDS

Fig. 1. Schematic representation of constructs for expression of HCV core and GFP linked to domain D2. Upper cartoons show the processing events in HCV polyproteins consisting of CE1E2 and CspmtE1E2. Black regions indicate signal peptides between the HCV structural proteins and the hatched segment represents the minimal sequence for domain D2 (amino acids 118-171). Black diamonds show the positions of SP cleavage sites that generate the HCV proteins and the white diamond denotes the SPP cleavage site in SP<sub>core-E1</sub>. SPP cleavage of CspmtE1E2 is blocked by mutations at amino acid positions 180, 183 and 184 (ASC to VLV) in SP<sub>core-E1</sub>. The lower diagram shows a cartoon of the GFP/D2 chimeric protein and the D2 amino acid sequence from HCV strain Glasgow.

Fig. 2. Targeting of GFP/D2 to LDs. *A*, Subcellular localization of LDs and GFP/D2 (i-iii) and GFP (iv-vi) in HuH-7 cells by confocal microscopy. Images were recorded after staining with BODIPY<sub>558/568</sub> C<sub>12</sub> to label LDs and fixation with paraformaldehyde. The scale bar in panel i represents 5 $\mu$ m. *B*, Co-fractionation of GFP/D2 with LDs in sucrose density gradients. Cells expressing either GFP or GFP/D2 were lysed and extracts were subjected to centrifugation on sucrose gradients. Fractions were removed from the top (1) to the bottom (11) of gradients, separated by SDS-PAGE, and examined by Western blot analysis using antibodies against GFP, calnexin and ADRP.

Fig. 3. Far UV CD analyses of synthetic peptides C[118-149] and C[141-171] in membrane mimetic environments. CD spectra of peptides C[118-149] (*A*, *B*) and C[141-171] (*C*, *D*) were recorded in 5 mM sodium phosphate buffer, pH 7.5, (H<sub>2</sub>O solution, dashed line), complemented with either 50% TFE (solid line) or the following detergents; 100 mM sodium dodecyl sulfate (SDS, dotted line) and 100 mM n-dodecyl  $\beta$ -D-maltoside (DM, large dashed line) in panels *A* and *C*; 1% L- $\alpha$ -lysophosphatidylcholine (LPC, dashed line) and 100 mM dodecyl phosphocholine (DPC, dotted line) in panels *B* and *D*. To facilitate data comparisons in different panels, the spectra of C[118-149] and of C[141-171] in 50% TFE were used for reference purposes in panels *A*, *B* and panels *C*, *D* respectively.

Fig. 4. NMR data analysis of peptides C[118-149] and C[141-171]. *A*, Summary of sequential (i, i+1), and medium range (i, i+2 to i, i+4) NOEs for each residue. NOE intensities are indicated by height of the bars. Asterisks indicate that the presence of a NOE is not confirmed because of overlapping resonances. The sequence 141-149 present in both peptides is underlined. *B*, <sup>1</sup>H $\alpha$  chemical shift differences (in ppm) for each residue. Differences were calculated by subtraction of the experimental values to the random coil

conformation values (54). The dotted line indicates the standard threshold value for an  $\alpha$ -helix ( $\Delta H\alpha$  of -0.1 ppm; 37). Boxed amino acids correspond to residues located in  $\alpha$ -helical regions as deduced from the NMR data. *C*, Helix projections of  $\alpha$ -helices 119-136 (left) and 148-164 (right) showing their amphipathic nature. Outline and boldface letters correspond to polar and hydrophobic residues, respectively.

**Fig. 5.** Intracellular localization of GFP/D2 mutants and LDs. *A*, The schematic shows the D2 amino acid sequence aligned with the limits of Helix I, HL and Helix II. Residues that were mutated to either glutamic (E) or aspartic acid (D) are indicated. *B*, Schematic showing the limits of the D2 coding sequences required for targeting of GFP to LDs. Association with LDs was assessed by confocal microscopy and the percentage of cells with GFP fluorescence associated with LDs is shown. “\*” indicates cells with very weak fluorescence at the surface of LDs. *C*, Cells expressing mutant forms of GFP/D2 were incubated with BODIPY<sub>558/568</sub> C<sub>12</sub> to stain LDs. Confocal microscopy images were recorded after fixation with paraformaldehyde. The GFP proteins expressed are indicated for each panel and partial co-localisation with LDs is arrowed. Scale bar represents 5 $\mu$ m.

**Fig. 6.** Stability and localisation of HCV core protein with mutations in D2. *A*, Western blot analysis of wt and mutant HCV core proteins expressed in cells electroporated with the corresponding SFV/CE1E2 RNAs. Cells were incubated in the presence or absence of MG132 as indicated and extracts were analyzed by Western blot analysis using anti-core antibody R308. Stability of each mutant protein was estimated from the relative band intensities observed in the presence and absence of MG132. *B*, Cells expressing core mutants were incubated in the presence and absence of MG132, fixed with methanol, and analyzed by indirect immunofluorescence using anti-core antibody R308 and anti-E2 antibody ALP98. Punctate nuclear localisation of core is arrowed in panel xi. Scale bar represents 5 $\mu$ m. *C*, Detection of core in extracts from cells electroporated with pSFV/C<sub>F130E</sub>E1E2 (lanes 1 and 2) and pSFV/C<sub>F130E/spmt</sub>E1E2 (lanes 3 and 4) by Western blot analysis. Cells were incubated either in the presence (lanes 2 and 4) or absence of MG132 (lanes 1 and 3). Extracts were probed with antisera R308.

**Fig. 7.** Summary of the LD-association characteristics and stability of D2 mutants. Small and hydrophilic amino acids are shown above the structural elements in D2 and hydrophobic residues are below the structures. Mutations are indicated as having either no appreciable effect (+) or weakened binding (-) of GFP/D2 to LDs. Core mutants were assessed as either stable (+) or having reduced (<40%) stability (-) upon expression in the SFV system.

**Fig. 8.** Model depicting the role of D2 in the folding of HCV core and LD targeting. (i) The HCV polyprotein is translocated to the ER membrane and cleaved by SP to generate the N-terminus of E1. (ii) Interaction of D2 with the ER membrane induces folding of Helix I and II, which promotes folding of D1. (iii) and (iv) SPP cleavage within SP<sub>core-E1</sub> releases the mature form of core that remains attached to the ER membrane via the structural motifs in D2. Alternatively, after SPP cleavage, incorrectly folded protein is targeted to the proteasome for degradation. (v) Mature core protein transfers from the ER membrane to the surface of LDs that are located at the ER membrane.

Table I: Association efficiency of GFP/D2 mutants to LDs

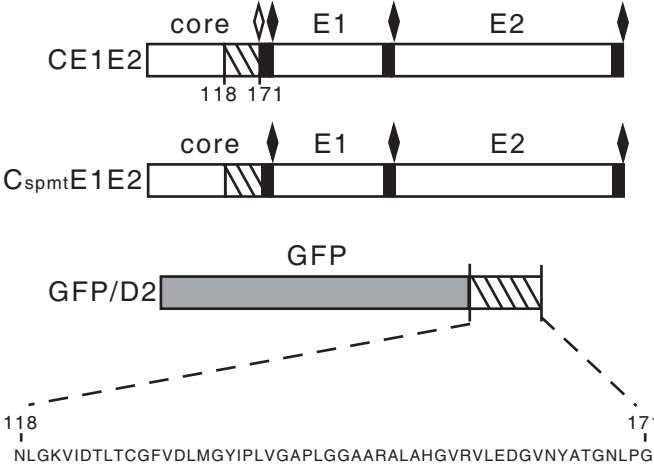
Construct	Class A <sup>a</sup>	Class B <sup>b</sup>
wt	100	0
L119E	88	12
K121E	89	11
I123E	90	10
L126E	50*	50
C128E	95	5
F130E	0	100
L133E	20*	80
M134D	75*	25
L139D	77	13
V140D	30*	70
L144E	30*	70
G145E	100	0
A147E	50*	50
R149E	100	0
A150E	90*	10
L151E	40*	60
L158E	100	0

<sup>a</sup> Percentage of cells with GFP fluorescence at LDs

<sup>b</sup> Percentage of cells with only diffuse GFP fluorescence

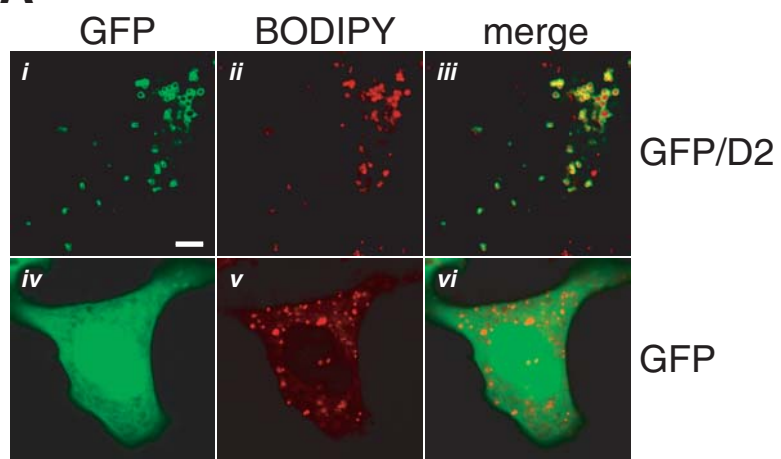
\* Percentage of cells with very weak LD association

**Fig. 1**

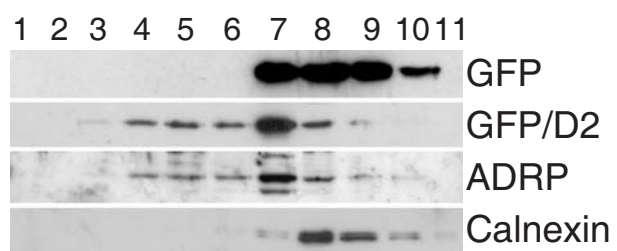


**Fig. 2**

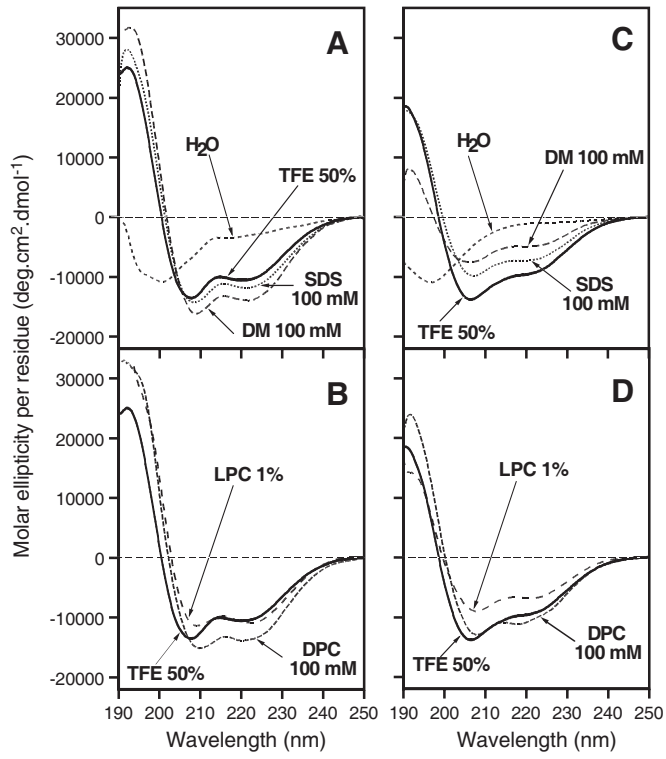
**A**



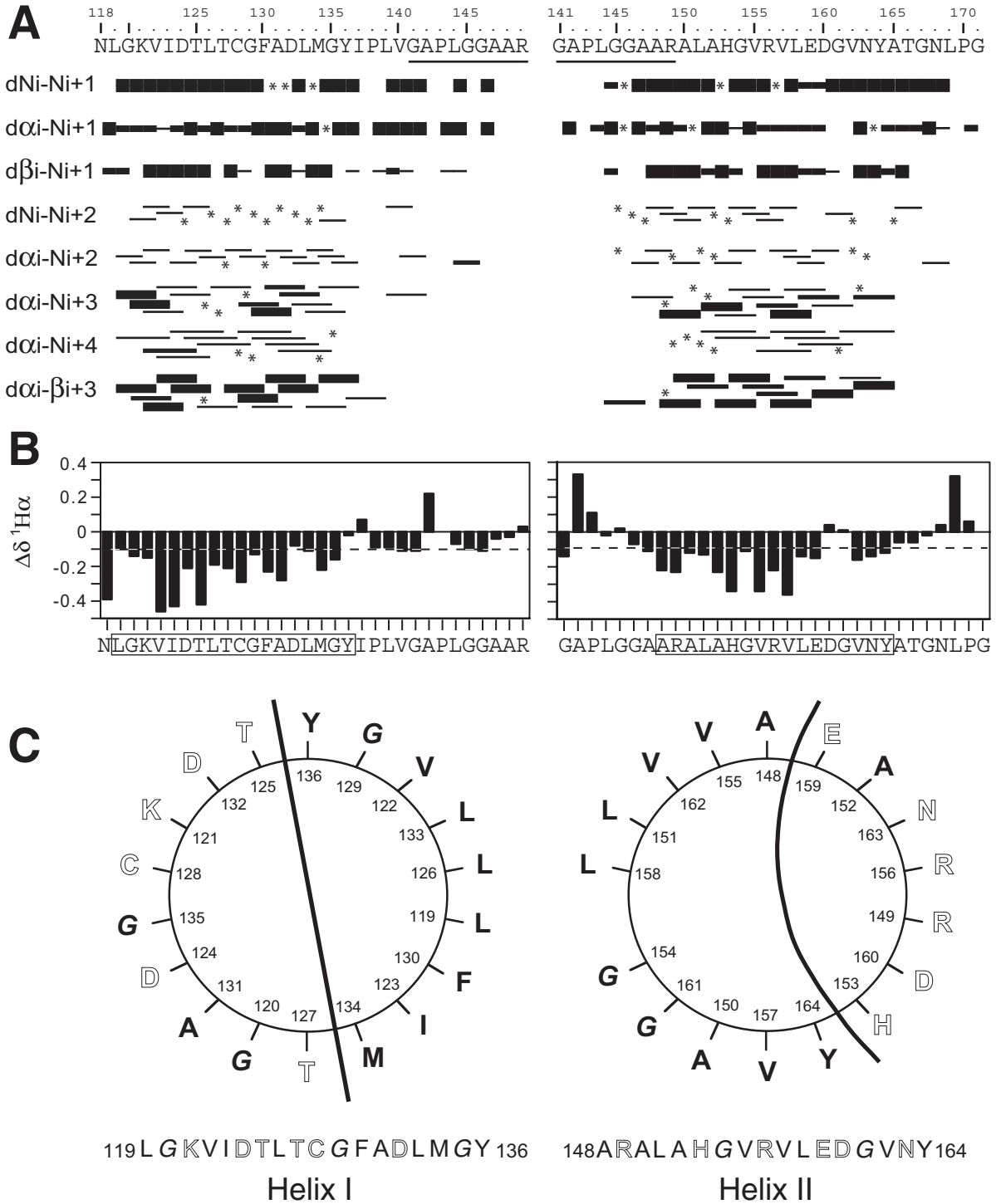
**B**



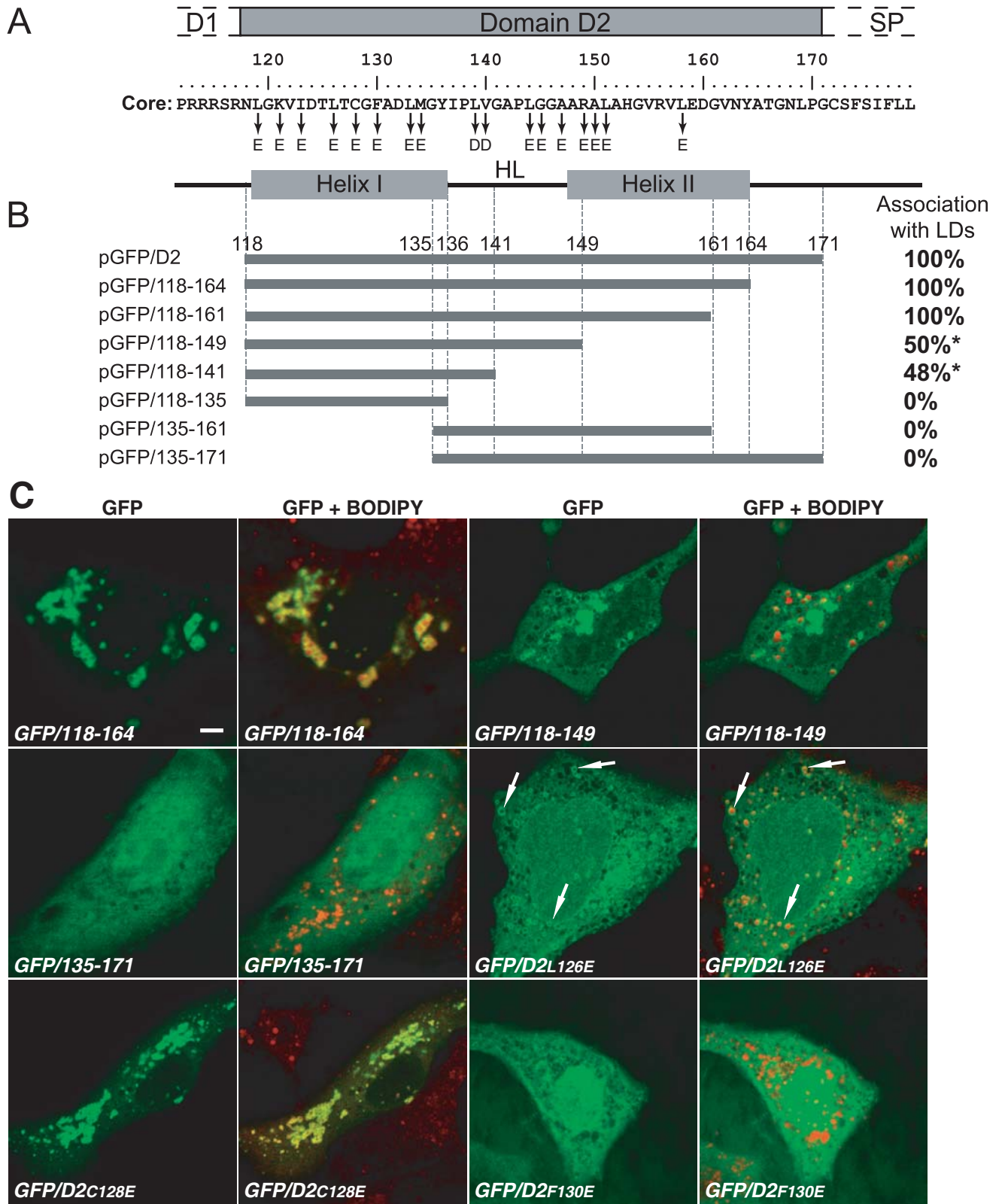
**Fig. 3**



**Fig. 4**

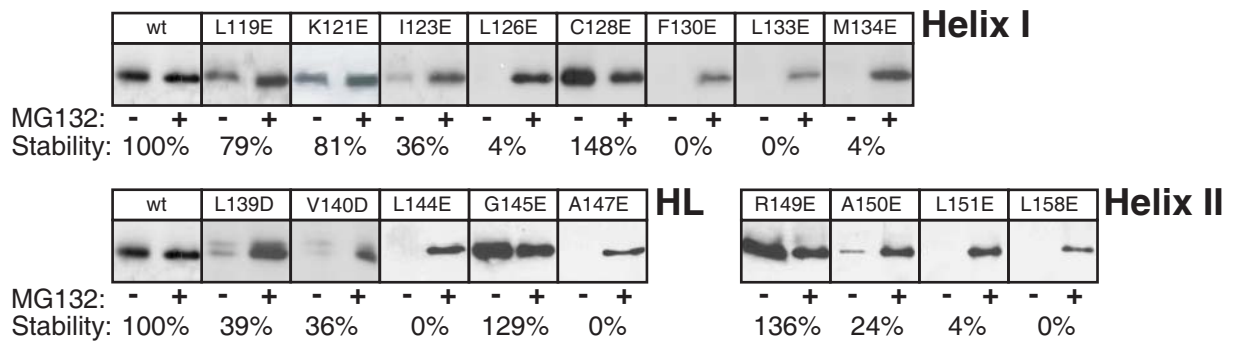


**Fig. 5**

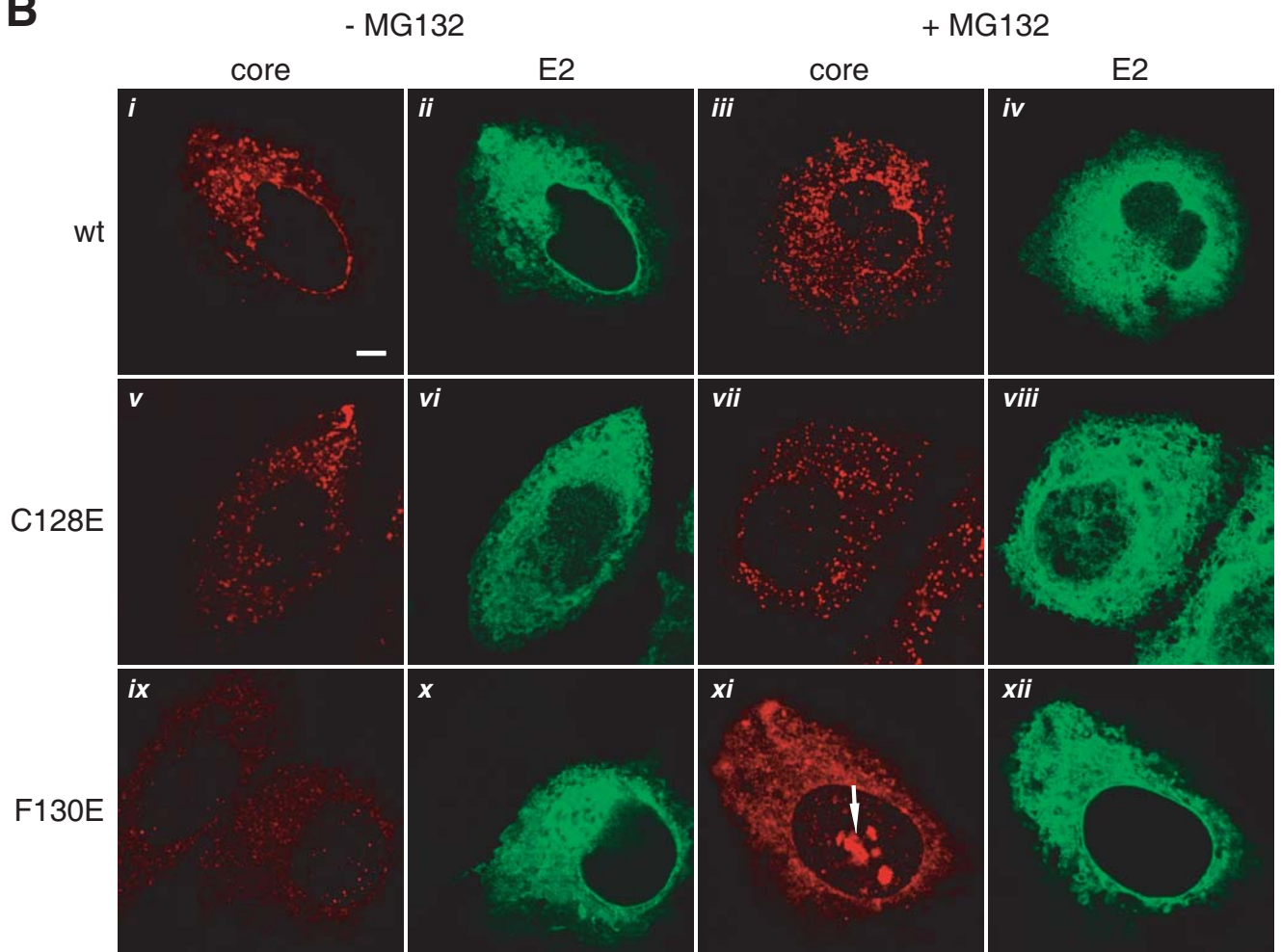


**Fig. 6**

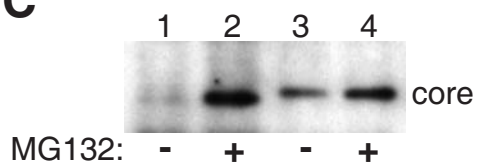
**A**



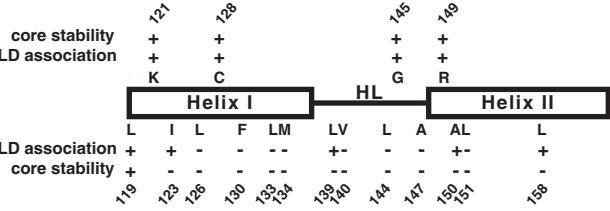
**B**



**C**



**Fig. 7**



**Fig. 8**

