Structural basis of the oligomerization of hepatitis delta antigen

Harmon J Zuccola[1], James E Rozzelle[2], Stanley M Lemon[3], Bruce W Erickson[2] and James M Hogle[1,4]*

Background: The hepatitis D virus (HDV) is a small satellite virus of hepatitis B virus (HBV). Coinfection with HBV and HDV causes severe liver disease in humans. The small 195 amino-acid form of the hepatitis delta antigen (HDAg) functions as a *trans* activator of HDV replication. A larger form of the protein containing a 19 amino acid C-terminal extension inhibits viral replication. Both of these functions are mediated in part by a stretch of amino acids predicted to form a coiled coil (residues 13–48) that is common to both forms. It is believed that HDAg forms dimers and higher ordered structures through this coiled-coil region.

Results: The high-resolution crystal structure of a synthetic peptide corresponding to residues 12 to 60 of HDAg has been solved. The peptide forms an antiparallel coiled coil, with hydrophobic residues near the termini of each peptide forming an extensive hydrophobic core with residues C-terminal to the coiled-coil domain in the dimer protein. The structure shows how HDAg forms dimers, but also shows the dimers forming an octamer that forms a 50 Å ring lined with basic sidechains. This is confirmed by cross-linking studies of full-length recombinant small HDAg.

Conclusions: HDAg dimerizes through an antiparallel coiled coil. Dimers then associate further to form octamers through residues in the coiled-coil domain and residues C-terminal to this region. Our findings suggest that the structure of HDAg represents a previously unseen organization of a nucleocapsid protein and raise the possibility that the N terminus may play a role in binding the viral RNA.

Addresses: [1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA, [2]Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, [3]Department of Microbiology and Immunology, The University of Texas Medical Branch at Galveston, Galveston, TX 77555-1019, USA and [4]Committee on Higher Degrees in Biophysics, Harvard University, 240 Longwood Avenue, Boston, MA 02115, USA.

*Corresponding author.
E-mail: hogle@hogles.med.harvard.edu

Key words: coiled coil, delta antigen, hepatitis D, leucine zipper

Received: **23 March 1998**
Revisions requested: **23 April 1998**
Revisions received: **6 May 1998**
Accepted: **6 May 1998**

Structure 15 July 1998, **6**:821–830
http://biomednet.com/elecref/0969212600600821

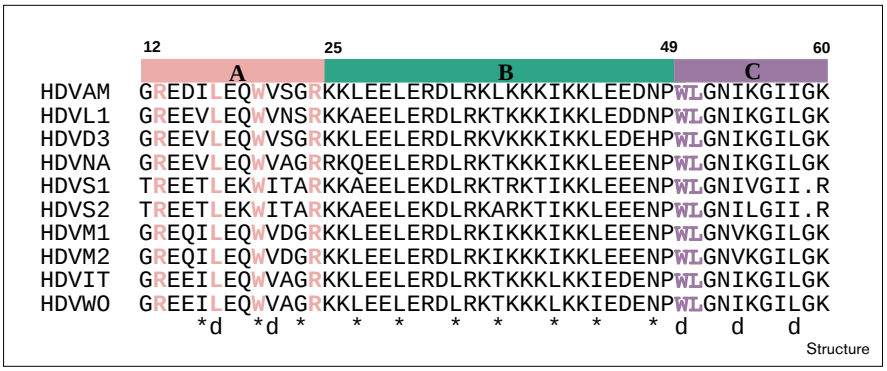
Introduction

Coinfection of hepatitis D virus (HDV) with hepatitis B virus (HBV) causes severe, and often fatal, liver disease in humans, and is the most common cause of fulminant viral hepatitis [1]. The virus is a subviral satellite of HBV, requiring the hepatitis B surface antigen (HBsAg) for assembly and cell-to-cell transmission [2]. The viral genome can replicate in the absence of HBV, however [3]. The viral genome is a 1.7 kilobase single-stranded circular RNA, which is approximately 70% complementary to itself [4] and forms a rod-like structure [5]. The virus is believed to replicate by a double rolling-circle mechanism in infected cells [6]. Both the genomic and antigenomic strands of the virus contain ribozymes [7–10] which are responsible for the reduction of multimeric viral genomes into unit length and for directing the religation of the linear genomes [11]. The antigenomic strand of the genome encodes the only viral protein known to be associated with HDV, the hepatitis delta antigen (HDAg) [4,12].

Early in the life cycle of the virus HDAg is expressed as a 195 amino-acid protein, the small hepatitis delta antigen (HDAg-S), which functions as a transactivator of HDV RNA replication [3]. Later in the life cycle of the virus, there is an RNA editing event that changes the UAG stop codon of the HDAg-S sequence to a UGG codon, encoding a tryptophan. This allows translation to proceed for an additional 19 amino acids, resulting in a 214 amino-acid form of the protein, the large delta antigen (HDAg-L). HDAg-L is a potent inhibitor of HDV replication [13,14] and is also involved in packaging the viral RNA [15–17] and the copackaging of the small antigen into the viral particle [15,17,18]. Both the large and small antigens are highly specific RNA-binding phosphoproteins [19,20], and have been shown to recognize specifically the viral rod-like structure of the HDV viral genomes [21]. Cross-linking studies have shown that both proteins can exist as either homomultimers (all small antigen or all large antigen) or as heteromultimeric structures (a mixture of small and large antigen) [22–24].

There have been a number of structure–function studies of both the large and small delta antigens. The N-terminal one third of HDAg contains a putative coiled-coil sequence [22–24], which is followed by a bipartite nuclear localization signal [25]. The middle portion of HDAg contains two arginine-rich motifs that have been shown to bind to the viral RNA [26]. The C-terminal segment of HDAg-S is proline- and glycine-rich [27]. HDAg-L is prenylated at the extreme C terminus [28,29], and it is believed that this part of the

Figure 1



Sequence alignment of ten serotypes of HDAg between amino acids 12 and 60. Residues marked with an asterisk make up the a and d positions in the heptad repeat in the predicted coiled-coil region. Bold residues in pink and purple are involved in the hydrophobic interactions in the dimer between the two termini. The residues marked with 'd' are involved in the dimer–dimer interface. A (pink), B (green) and C (purple) represent the sequence segments described in the text.

molecule interacts with HBsAg and the membranes of the endoplasmic reticulum [30,31]. There is also some evidence that common segments of the large and small antigens may have subtly different conformations [32,33].

The coiled-coil domain has been shown to be required for a number of the functions of both small and large delta antigens. Mutations that destroy or alter the coiled-coil domain either greatly reduce or totally eliminate the ability of HDAg-S to function as a *trans* activator of replication [18,22]. These same mutations also prevent HDAg-L from inhibiting HDV RNA replication and inhibit its function in mediating the copackaging of HDAg-S [18,22]. Transfection of cells undergoing HDV replication with a plasmid containing just the N-terminal one-third of HDAg (which contains the coiled-coil domain) inhibited HDV replication [22]. But removal of the coiled-coil domain does not prevent HDAg from binding the viral RNA [20], nor does it prevent HDAg-L from packaging the viral RNA [34]. A 'black sheep' model has been proposed for the mechanism of inhibition of HDV replication. HDAg-L is believed to disrupt the homo-oligomeric HDAg-S multimers, essentially 'poisoning' the HDAg-S complex [22]. Although the precise role of HDAg-S in replication of HDV is unknown, the protein is not a polymerase: HDV RNA amplification is thought to be mediated by host-cell RNA polymerase II [35,36].

Biophysical studies were undertaken to examine the coiled-coil domain of HDAg [37]. A peptide was synthesized that corresponds to residues 12–60 of HDAg, referred to here as δ12–60(Y). The peptide also included a C-terminal tyrosine, so that the peptide could be labeled with I^{125} for use in a radioimmunoassay. The peptide sequence was conceptually divided into three segments based on the presence of two potential helix breakers, Gly23 and Pro49: segments A (residues 12–24), B (residues 25–49), and C (residues 50–60); see Figure 1. The full-length peptide δ12–60(Y) and two shorter peptides that corresponded to regions A+B and B+C were synthesized. A number of biophysical experiments, including circular dichroism (CD), mass spectrometry, and analytical ultracentrifugation, clearly showed that the δ12–60(Y) peptide was largely helical and formed a coiled coil [37]. The shorter peptides formed much less stable structures and were considerably less helical than δ12–60(Y). Human polyclonal antibodies from hemophilic patients who were chronic carriers of HBV and HDV reacted with the δ12–60(Y) peptide in both an ELISA and a sandwich radioimmunoassay [37,38]. Subsequent studies indicated that monoclonal antibodies against the peptide recognized a conformational epitope only presented by the full-length peptide and not the shorter, extensively overlapping peptides [37].

We describe here the crystal structure of the peptide δ12–60(Y) to 1.8 Å resolution. The structure of the peptide lends new insights into the mechanism by which HDAg dimerizes and further associates into higher ordered structures. The structure also explains why residues C-terminal to the predicted coiled-coil domain, and the helix-breaking proline residues are important for the stabilization of the coiled-coil structure. The peptide structure has important consequences for the oligomerization of HDAg *in vivo*. The unique octameric structure observed in the crystal structure also suggests that the N terminus of the molecule may have a previously undetermined function.

Results

Structure determination

Attempts to find a heavy atom derivative using the peptide with the wild-type sequence of the American strain of HDAg failed. Thus, a new peptide was synthesized with a cysteine replacing Ser22 (this residue demonstrates considerable variation among different strains of HDV; Figure 1). The cysteine-mutant and wild-type peptides crystallized isomorphously. The presence of Cys22 allowed the preparation of a platinum terpyridine derivative, facilitating the determination of the structure using single isomorphous replacement with anomalous scattering (SIRAS)

Table 1

Data collection statistics.

	Native*	S22C Pt
Spacegroup	$P2_12_12$	$P2_12_12$
Unit cell (a, b, c)	109.2, 85.3, 29.4	110.3, 86.3, 29.6
Temperature of data collection (°C)	−160	−165
Resolution (Å)	15–1.73	86–2.5
Number of reflections	221,286	44,362
Number of unique reflections	28,279	10,013
Completeness† (%)	94 (35)	97
I/σ†	51 (7)	6.0
Multiplicity	7.8	4.4
R_{sym} †‡§ (%)	4.2 (18)	6.7 [14.0]
R_{iso} ¶ (%)	–	30.5
R_{cullis} #	–	0.62 (0.52)
$R_{cullis\ anom}$ ¥	–	0.84
Phasing power**	–	2.2 (1.7)

*Data are from two crystals. †Numbers in parentheses represent values in the highest resolution shell. ‡$R_{sym} = \Sigma_{(h,k,l)} | I_{(h,k,l)} - \langle I_{(h,k,l)} \rangle | / \Sigma_{(h,k,l)} \langle I_{(h,k,l)} \rangle$, where $\langle I_{(h,k,l)} \rangle$ represents the sigma-weighted average intensity of symmetry-equivalent reflections. §The number in square brackets represents $R_{anom} = \Sigma | \langle I+_{(h,k,l)} \rangle - \langle I-_{(h,k,l)} \rangle | / \Sigma (\langle I+_{(h,k,l)} \rangle + \langle I-_{(h,k,l)} \rangle)$, where $\langle I+/-_{(h,k,l)} \rangle$ represents the statistically weighted average intensity of symmetry-equivalent reflections. ¶$R_{iso} = \Sigma_{(h,k,l)} | (F_{PH} - F_P) | / \Sigma(F_P)$. #$R_{cullis} = \Sigma_{(h,k,l)} | |F_{PH}| - |F_P + F_H| | / \Sigma_{(h,k,l)} |F_{PH} - F_P|$; number in parentheses represent R_{cullis} for centric reflections. ¥$R_{cullis\ anom} = \Sigma_{(h,k,l)} | |F_{PH+} - F_{PH-}|_{obsvd} - |F_{PH+} - F_{PH-}|_{calc} / \Sigma_{(h,k,l)} |F_{PH+} - F_{PH-}|_{obsvd}$. **Phasing power $= \langle |F_H| / ||F_{PH}| - |F_P + F_H|| \rangle$; the number in parentheses is the phasing power for centric reflections.

Table 2

Refinement statistics.

Resolution range (Å)	15–1.8
$R_{working}$* (%)	22.5
R_{free}† (%)	27.0
Non-hydrogen protein atoms	1785
Solvent atoms	114
Rms from ideal geometry	
bond lengths (Å)	0.007
bond angles (°)	1.0
dihedral angles (°)	16.9
impropers (°)	0.59
Average B factor overall (Å²)	29.3
mainchain	22.5
sidechain	32.4
solvent	34.7

*$R_{working} = \Sigma_{(h,k,l)} | (|F_o| - |F_c|) | / \Sigma(F_o)$, for a working set composed of 90% of the data. †$R_{free} = \Sigma_{(h,k,l)} | (|F_o| - |F_c|) | / \Sigma(F_o)$, for a test set composed of 10% of the data selected randomly.

methods (Table 1). Retrospective examination of the model confirmed that the Pt was indeed bound to the sulfur of Cys22.

The solvent-flattened map was easily interpretable, and clearly showed two dimers in the asymmetric unit. Rounds of positional refinement, simulated annealing, temperature factor refinement using X-PLOR [39], and manual rebuilding using O [40] led to the current model (Table 2, Figure 2). The current model has an R factor of 22.5% and a free R factor of 27% with good geometry (root mean square deviation [rmsd] bond 0.007 Å and rmsd bond angles 1.0°). A number of sidechains exposed to the large solvent channel, as well as the first residue in the chain and the last residue in one of the chains, are disordered. The four monomers in the asymmetric unit superimpose well onto one another, with an average rmsd for mainchain atoms of 0.81 Å and for all non-hydrogen atoms 1.51 Å. The main differences in the monomers are those residues involved in crystal packing interactions.

Structure description

Each monomer is composed of a long, N-terminal helix, approximately 60 Å in length, interrupted by a sharp bend at Pro49, and continuing on into another short helix. The long helices of each of two monomers wrap around each other forming an antiparallel coiled coil (Figure 3a,b), which straightens out at the N terminus. Only one of the four possible salt bridges between Glu31 and Lys38 is seen. In the other three cases, the charged groups are slightly farther apart (3.8 Å, 4.2 Å and 4.4 Å) and the sidechains are hydrogen-bonded to nearby solvent molecules. The

Figure 2

The final atomic model superimposed upon a portion of the final 1.8 Å resolution X-PLOR $2F_o - F_c$ map. The map is contoured at 1.2σ and shows the residues involved in the interaction between monomers at the A and C regions. The orientation is similar to that in Figure 4. (The figure was produced using the program BOBSCRIPT.)

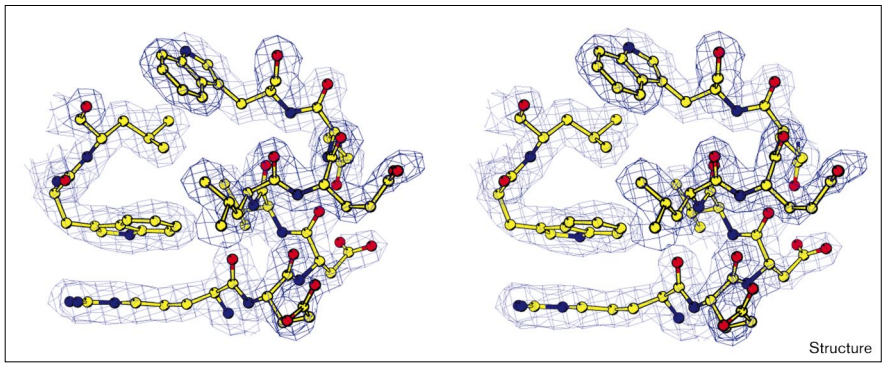
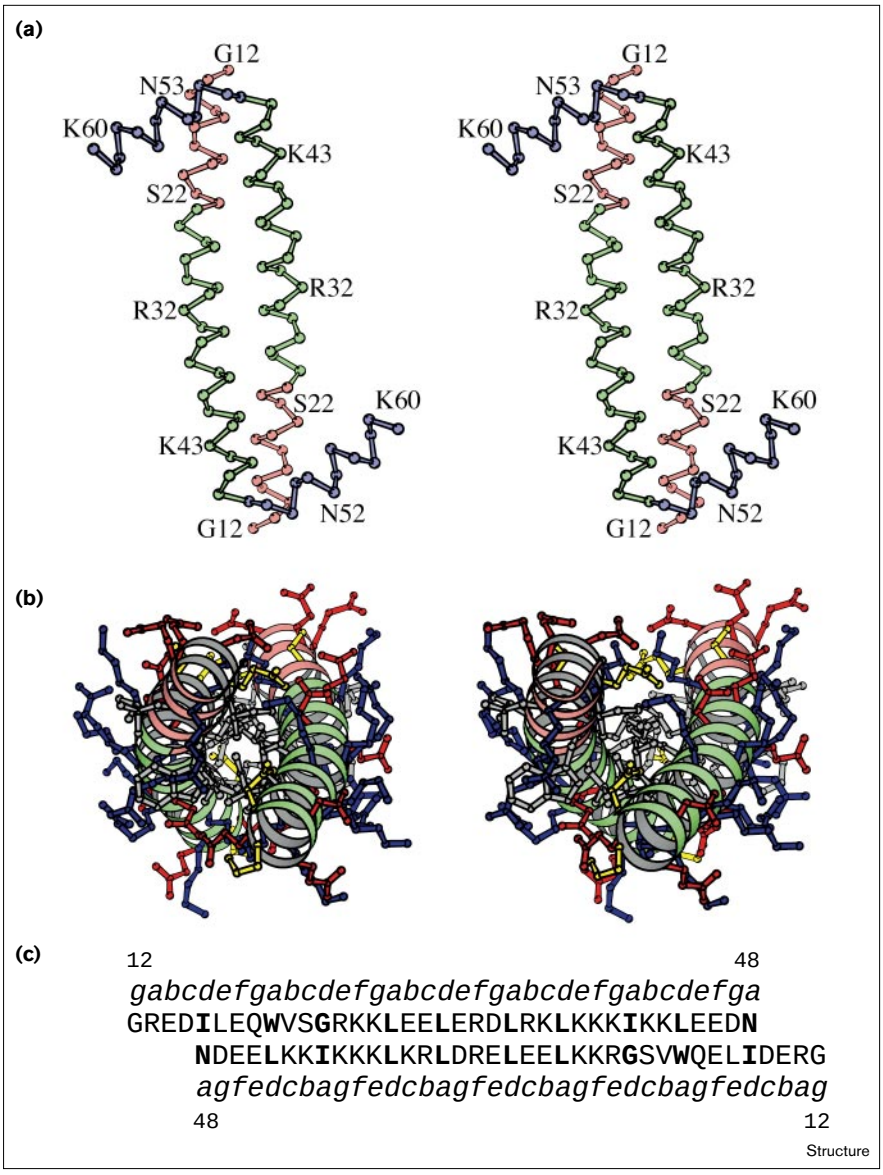


Figure 3

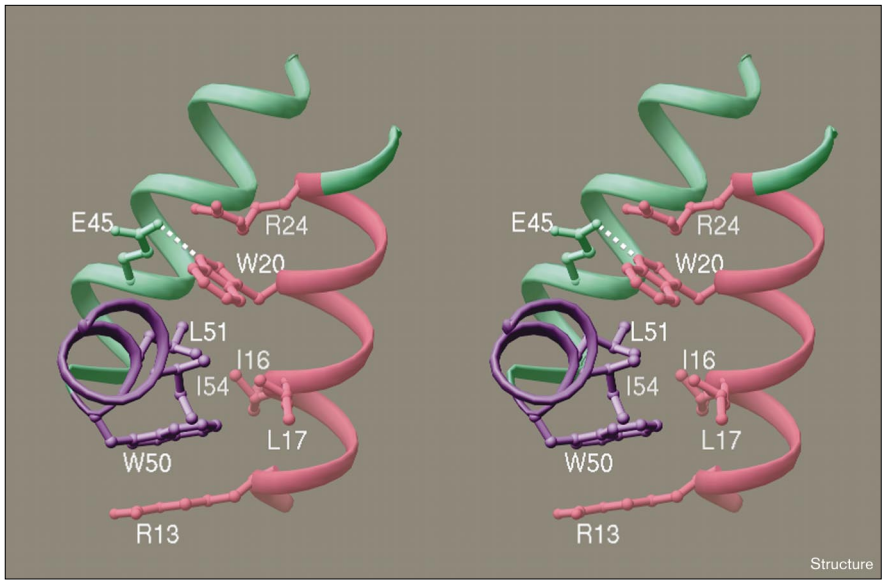


Overall organization of the HDAg dimer. **(a)** Cα trace of the peptide δ12–60(Y). The three regions of the peptide are color-coded: A pink; B green; and C purple. The two monomers form an antiparallel coiled coil. The individual helix takes a sharp bend at Pro49. **(b)** Ribbon diagram of the view in (a) rotated 90° along the horizontal axis. The sidechains have been added and the C region of the peptide (residues 50–60(Y)) has been removed for clarity. Sidechains are color-coded: hydrophobic, gray; polar, yellow; acidic, red; and basic, blue. Although the majority of the hydrophobic sidechains are packed in the interior of the coiled coil, Trp20 is flipped out of the core of the long helix. The figure also illustrates that the two helices clearly wrap around one another. **(c)** The amino-acid sequence of the long helix formed from residues 12–48 displayed in the antiparallel orientation of the peptide. The letters above the amino-acid sequence represent the heptad repeat $(abcdefg)_n$ where the *a* and *d* residues tend to be hydrophobic. Residues involved in the heptad repeat at the *a* and *d* positions are shown in bold.

sidechain of Glu45 is hydrogen-bonded to the indole nitrogen of Trp20. The sidechain of Asn48, which is located at the C terminus of the long helix, completes the hydrogen-bonding pattern of the helix by making a hydrogen bond back to the mainchain oxygen of Leu44. The formation of the dimer buries 2650 Å^2 of surface area, approximately 26% of the total surface area.

Although the majority of residues in the heptad repeat (Figure 3c) of the predicted coiled-coil region do pack as expected, Trp20 does not. Instead, the sidechain of Trp20 is flipped away from the core of the coiled coil and into a hydrophobic region formed between segment A (residues 12–24) of one monomer, and segment C (residues 50–60) of its partner within the peptide dimer. The dimer shows primarily hydrophobic interactions between residues in the A and C regions. Ile16, Leu17, Trp20, Trp50, and Leu51 are the sidechains primarily involved in this hydrophobic region, which is capped by the aliphatic portion of the sidechains of Arg13 and Arg24 (Figure 4). The primary non-hydrophobic, monomer–monomer interactions near this region involve the formation of a hydrogen bond between Trp20 and Glu45 (Figure 4). The heptad repeat is also unusual in that it contains a glycine at position 23. If the monomers were oriented in a parallel fashion, a large 'hole' in the middle of the hydrophobic core of the dimer would result. But as the strands are antiparallel, the large sidechain of Ile41 packs into the hole formed by Gly23.

Figure 4



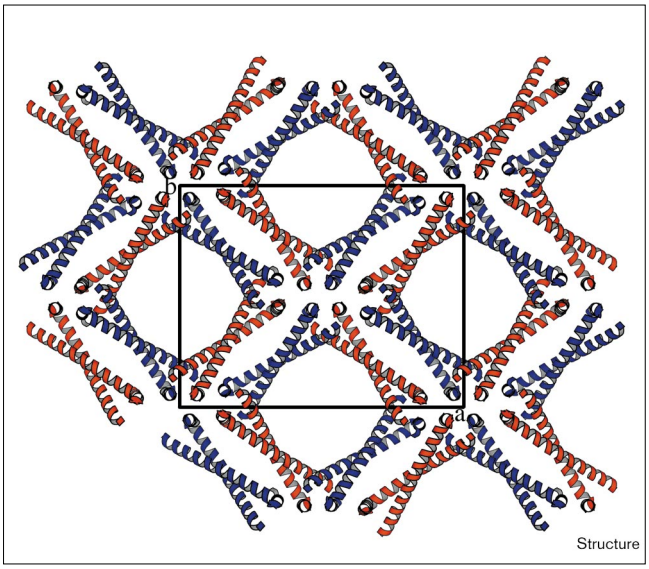
The dimer is stabilized by hydrophobic interactions other than those provided by residues in the heptad repeat. Residues from the N termini of each monomer, Ile16, Leu17 and Trp20 from one monomer and Trp50, Leu51 and Ile54 from the other, form a hydrophobic core which is protected from solvent by the aliphatic portions of Arg13 and Arg24. There is also a hydrogen bond between the sidechain of Glu45 and the indole nitrogen of Trp20 (O–N distance 2.8 Å). The three regions of the peptide are color-coded: A pink; B green; and C purple. (This figure was produced using the program RIBBONS [58].)

In the crystal, each dimer associates with three other dimers to form a doughnut-like octamer (Figure 5). The octamer is widely open with a central 'hole', 50 Å in diameter. The open structure of the octamer is reminiscent of several other proteins, including that of proliferating cell nuclear antigen (PCNA), in which the hole that is formed is believed to encircle DNA [41]. It is this octameric structure which is the translational repeating unit in the crystal (Figure 5). The dimer–dimer interface is a four-helix bundle formed across the crystallographic twofold axis. The interface of the two dimers consists of hydrophobic residues in region A of the coiled-coil domain (Leu17 and Val21), but also includes residues C-terminal to the coiled-coil domain, region C, between residues 50 and 60 (Trp50, Ile54, Ile57 and Ile58; Figure 6), essentially extending the hydrophobic core mentioned above. Trp50 is involved in the formation of both the dimer and the octamer. Formation of the octamer buries an additional 800 Å^2 of surface area per monomer, which means that approximately 40% of the total surface area of each monomer is buried. The 50 Å diameter hole framed by the four dimers is lined with basic sidechains (Figure 7).

Mass spectrometry

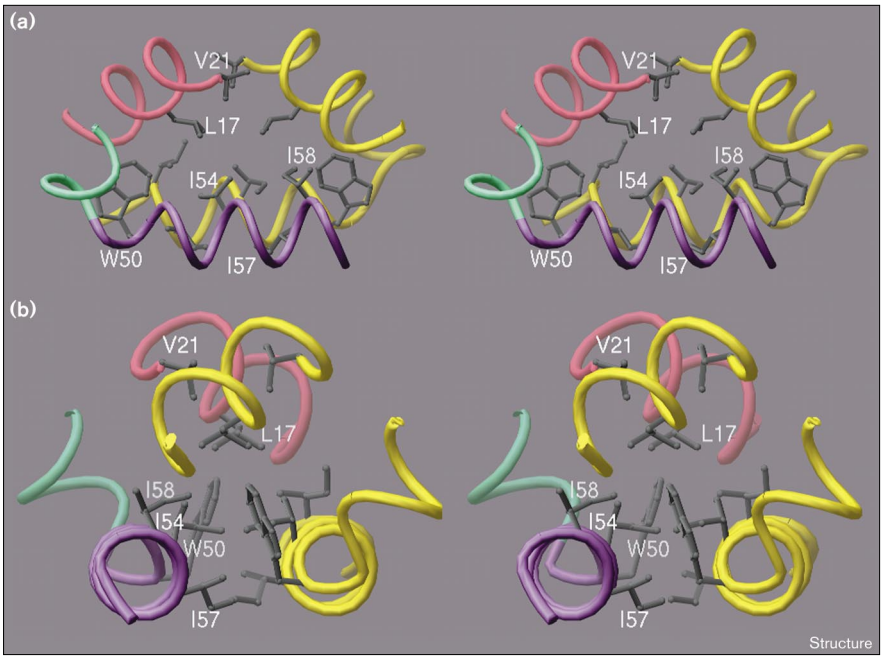
Previous studies have suggested that both the peptide and natural HDAg derived from infected liver form multimers in solution [23,37]. In order to investigate the significance of the octamer formed by the peptide, MALDI-TOF mass spectrometry was used to determine the mass of monomeric and oligomeric forms of recombinant small delta antigen, rHDAg-S. The uncross-linked protein has a mass of 21,832 Da (Figure 8a), which is the correct mass (within 0.01%) of the amino-acid sequence of the American strain of HDAg-S (Genbank accession number M28267) minus the first methionine residue. The primary species of the cross-linked rHDAg-S had a mass of 176,282 Da (Figure 8b). Indeed, the M+1 and

Figure 5



The interactions of dimers in the $P2_12_12$ unit cell. The unit cell is outlined in black and the directions of the a and b axes are shown. Four dimers come together to form an octameric complex forming a pseudo-centered (C222) cell. The two independent copies of the dimer in the asymmetric unit are colored orange and blue. The view is looking down the crystallographic twofold axis. (This figure was produced using the program MOLSCRIPT [59].)

Figure 6



The dimer–dimer interface extends the hydrophobic core formed by the ends of the two monomers. **(a)** The dimer–dimer interface is composed of a four-helix bundle made from the N and C termini of two dimers, one from across the crystallographic twofold axis. Hydrophobic residues from the four helices pack in the interface, Leu17, Val21, Trp50, Ile54, Ile57 and Ile58. For clarity, one (unlabeled) dimer has been colored yellow and the other (labeled) dimer is colored according to the scheme used in Figure 1. **(b)** The view in (a) rotated 90° around the y axis. (This figure was created with the program RIBBONS [58].)

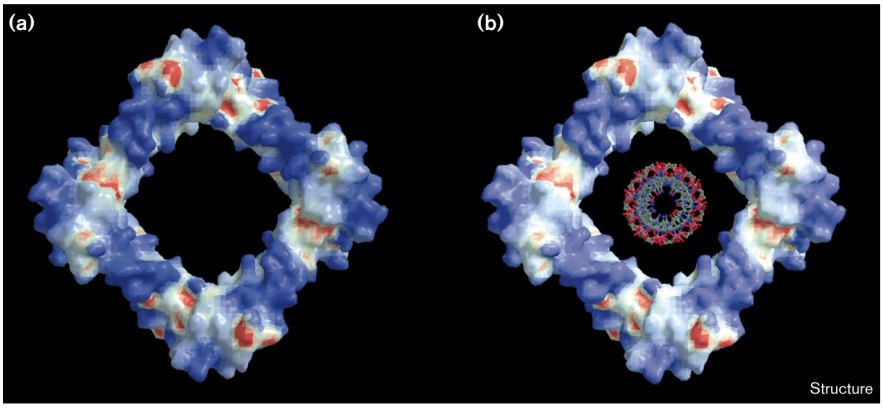
M+2 peaks of the octamer were the only significant peaks in the spectrum. The ratio of the masses of the cross-linked species to the monomer is 8.1:1.

Discussion

Oligomerization of HDAg

The structure of the δ12–60(Y) peptide raises a number of interesting questions about the function of the coiled-coil domain of HDAg. When the HDAg open reading frame was originally examined, amino acids from residue 13 to 47 were identified as possibly forming a coiled coil. Glutaraldehyde cross-linking studies of full-length HDAg, as well as of the peptide, confirmed the formation of dimers, tetramers and higher-ordered structures [23,37]. The crystal structure of the peptide clearly shows how monomers come together to form antiparallel dimers as well as a higher-ordered octameric structure. The structure of δ12–60(Y) also agrees well with previous CD studies of the peptide, which indicated that the two ends of the peptide (regions A and C) were important for the structural stability of the coiled coil [37]. Shorter synthesized peptides that were missing either the A or C regions (A+B and B+C) were significantly less helical than the full-length peptide (A+B+C; 37%, 45% and 84% respectively at 37°C). The peptide structure shows that hydrophobic residues from the N terminus of one monomer (region A), not involved in the heptad repeat, interact with residues outside of the predicted coiled-coil domain

Figure 7

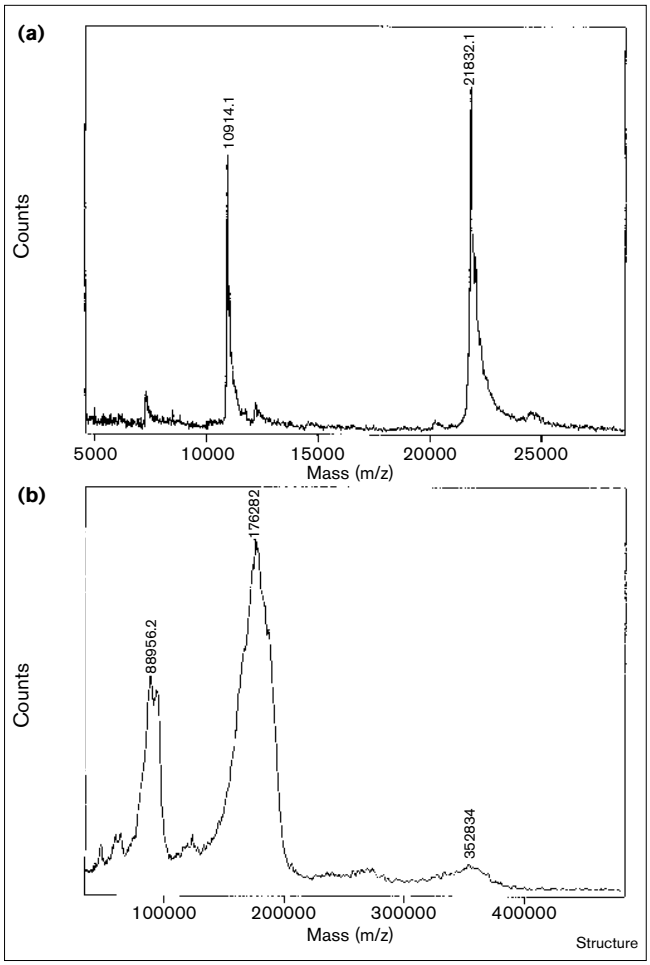


Solvent-accessible surface showing electrostatic potential of the HDAg octamer. **(a)** A GRASP electrostatic potential surface of the octameric δ12–60(Y) peptide contoured at = 10 kT/e (positive potential blue) and = 10 kT/e (negative potential red). Residues Lys26 and Lys38, which had been modeled as alanine, were changed to lysine for this calculation. The edges and the lining of the large 50 Å hole are basic. **(b)** The hole formed by the octamer is large enough to accommodate an RNA molecule. (The electrostatic surface was calculated using the program GRASP [60] and rendered using RASTER3D [61].)

near the C terminus of the other monomer (region C) to form a hydrophobic core (Trp20, Leu24, Trp50, Leu51) sandwiched between Arg13 and Arg24. This may stabilize the structure by keeping the ends of the helix from fraying. An additional stabilizing feature is a hydrogen bond between the sidechain of Glu45 and the indole nitrogen of Trp20. These hydrophobic residues, as well as the glutamic acid residue, are highly conserved in the 11 different strains of HDV identified to date (Figure 1). In fact, they are more conserved than those residues in the heptad repeat making up the hydrophobic core of the long helix (Figure 1).

The observation that rHDAg-S forms octamers means that the octamer form seen in the crystal may not be an artifact of crystallization, but rather may represent the true state of the oligomerization of the delta antigen. A study by Chang and colleagues found that a deletion in HDAg-L, just C-terminal to the coiled-coil domain (residues 50–75), prevented HDAg-S from being copackaged with HDAg-L [42]. HDAg-L with this same deletion could not inhibit HDV replication, whereas a deletion in HDAg-L of residues 65 to 75 could. This suggested that the coiled-coil domain alone is not sufficient for the interaction between the large and small antigens, and that a subdomain between residues 50 and 65 is also necessary for this interaction. The crystal structure of δ12–60(Y) indeed shows the importance of residues between 50 and 60 in the formation of the peptide oligomer. They are not only involved in stabilizing the δ12–60(Y) dimer (Trp50 and Leu51), but are also involved in the formation of the dimer–dimer interface (Trp50, Ile54, and Ile58).

Figure 8



MALDI-TOF analysis of **(a)** rHDAg-S and **(b)** the glutaraldehyde cross-linked protein. The protein was prepared as described in the Materials and methods section.

Black sheep model revisited

Before the studies described here, the overall organization of the HDAg oligomer was unknown. Lai and coworkers [22], inferring from previous data that showed that as little as 12% of HDAg-L is needed to inhibit 90% of viral activity [13], proposed that as little as one part of HDAg-L in eight parts of HDAg-S could inhibit viral replication. Their black sheep model proposed that HDAg-L either disrupted the conformation of the oligomer of HDAg-S, therefore preventing it from binding to host factors, or that the presence of HDAg-L in the complex prevents the complex from interacting with host factors. This would seem in agreement with the peptide structure of octameric δ12–60(Y) and the results of the MALDI-TOF analysis.

If HDAg-L does disrupt the conformation of the oligomer of HDAg-S, it probably does not do so directly through the multimerization domain, given that the large and small delta antigens share the same sequence within this region. Rather, it is possible that this $\alpha_7\beta$ or $\alpha_n\beta_m$ structure can no longer interact with host factors. Also, as the C terminus of HDAg-L interacts with the endoplasmic reticulum (ER) membrane and with HBsAg for assembly, it could redirect the complex elsewhere in the cell, preventing the nuclear translocation of HDAg-S which is required for HDV replication.

Possible undetermined functions of the N terminus

The octamer that is formed by the peptide is reminiscent of proteins that form clamps around DNA, such as PCNA [41]. The 50 Å hole formed by the octameric structure is lined with basic sidechains, suggesting that the N terminus of the protein not only may act as a dimerization/oligomerization domain, but also that it may function either as a clamp around the viral RNA or other nucleic acid or perhaps even function as a spool for nucleic acid. There is a report that peptides corresponding to the extreme N-terminal portion of HDAg, residues 2 to 27 and 2 to 17, can bind the viral RNA [43,44]. As the δ12–60(Y) structure is missing residues 2 to 11, it is impossible to say what role they play in binding the viral RNA. Of the remaining

residues, only Lys25 and Lys26, which point into the hole of the octamer, seem likely to play a role in binding RNA by potentially binding the phosphate backbone of the viral RNA.

The large size of the hole may be necessary to accommodate the viral RNA which is only 70% self-complementary, and would possess a number of regions of bulged out single-stranded sequence, increasing the radius of gyration of the RNA as well as bending the RNA [45]. The octameric structure also implies that there may be as many as four RNA-binding domains on each side of the octamer, although the exact consequences of such an arrangement is presently unclear. It is possible that this portion of the molecule could also bind another protein, especially one that is acidic, such as the recently discovered delta antigen interacting protein A (dipA), a cellular protein of unknown function which has been found to interact with HDAg [46] and, based on its amino-acid sequence, would have an isoelectric point of 4.9.

A cautionary tale

Many investigators referred to the putative coiled-coil domain of the delta antigen as a leucine-zipper-like region. Experiments involving mutations in this region were interpreted assuming the coiled-coil domain of the delta antigen would resemble the parallel coiled coil of the bZIP family of transcription factors, such as GCN4. Along with the structure of the δ12–60(Y) peptide, there are other examples of molecules that dimerize through antiparallel coiled-coil domains, such as the *Escherichia coli* regulatory protein, AraC [47], and the replication terminator protein from *Bacillus subtilis* [48]. Although algorithms have been designed to determine the oligomerization state of a coiled coil [49,50], no program yet exists that attempts to determine the orientation of the predicted coiled coil. The discovery that this region forms an antiparallel coiled coil should be a reminder that without further biochemical or genetic evidence, a predicted coiled-coil domain may adopt either a parallel or antiparallel conformation.

Biological implications

The hepatitis delta antigen (HDAg), the sole protein made by the hepatitis delta virus (HDV), is essential for viral replication *in vivo*. Oligomerization of the protein is necessary for both the transactivating function of the small HDAg-S and the *trans* dominant inhibitory effect of the large HDAg-L. The structure of the peptide δ12–60(Y) that corresponds to the predicted coiled-coil domain of HDAg suggests that HDAg not only dimerizes through an antiparallel coiled coil but also forms octamers. Interestingly, the coiled coil is stabilized by hydrophobic residues C-terminal to the coiled-coil domain. These C-terminal residues interact with hydrophobic residues in the N terminus of the coiled-coil region. The hydrophobic core of the dimer is extended by further hydrophobic interactions at the interface between dimers in the octameric structure. In contrast to the rather promiscuous interactions between the coiled-coil domain, these unique interactions at the termini of the monomer and dimer interfaces might provide a good target for antivirals against HDV, as disruption of oligomerization can prevent replication *in vivo*.

The surprising octameric structure of the peptide suggests that the capsid of HDAg will look very different to the known structures of other viral nucleocapsid proteins. The octameric structure also suggests important implications for binding of HDAg to the viral RNA, because as many as four of the arginine-rich RNA-binding domains might be needed for binding to the viral RNA. The very basic hole in the octamer suggests that this portion of the molecule may act as a sort of 'clamp' around an acidic molecule; whether that molecule is the viral RNA, some other nucleic acid or a cellular factor has yet to be determined.

The exact function of HDAg in viral replication is still unclear. It has been suggested that the protein may only function as a shuttle, binding to the viral RNA and transporting it into the nucleus of the infected cell. It is possible that HDAg functions to recruit host cell transcriptional machinery to the viral RNA, but no such interaction has yet been identified. With the structure in hand, it should be possible to design experiments to further explore the possibility that the N terminus of the molecule has RNA-binding capabilities. A systematic examination of the amino acids involved in dimerization and oligomerization may allow the determination of the mechanism by which HDAg-L inhibits HDAg-S. Furthermore, the unique interactions at the termini of the coiled-coil region may provide a new framework to be exploited in the *de novo* design of stable antiparallel coiled coils.

Materials and methods

Crystallization and data collection

The δ12–60(Y) peptide was synthesized and purified as described previously [37]. The peptide was dissolved in 50 mM acetate, pH 4.8, 50 mM NaCl and brought to a concentration of 15 mg/ml. The crystals of the δ12–60(Y) peptide were grown at 22°C by the vapor diffusion method. Peptide (2 μl of a 15 mg/ml solution) was mixed with 2 μl of the reservoir solution containing 100 mM sodium acetate, pH 4.8, and 100 mM sodium citrate, pH 5.6, on a coverslip and then inverted over the reservoir solution. Crystals appeared within 3–4 days, and grew as large as $0.5 \times 0.3 \times 0.3$ mm. Crystals belonged to space group $P2_12_12$ with unit cell parameters a = 109.3 Å, b = 85.3 Å, c = 29.4 Å, $\alpha = \beta = \gamma = 90°$. When attempts to find a heavy atom derivative failed, a peptide was synthesized with Ser22 replaced by a cysteine, δS22C12–60(Y). The δS22C12–60(Y) peptide was reacted with an excess of platinum terpyridine, dialyzed overnight against water, and then freeze dried. The peptide was then reconstituted at 15 mg/ml in 50 mM acetate, 50 mM NaCl, 5 mM DTT, pH 4.8, and crystallized by the same conditions as that of the wild-type peptide. This peptide crystallized isomorphously with the δ12–60(Y).

The coverslips containing the crystals were inverted and cryosolvent (reservoir solution containing 30% glycerol) was slowly mixed with the drops and continuously replaced until no mixing was observed. The crystals were mounted in nylon loops and frozen directly in the nitrogen stream. Crystals used at Brookhaven were stored in liquid nitrogen until the time of data collection. Two native data sets were collected at Beamline X12C at the National Synchrotron Light Source at Brookhaven National Laboratory using X-rays of wavelength 1.15 Å (Table 1). The heavy atom data set was collected on a Siemens rotating anode with a multiwire detector (Table 1). Data from the native crystals were processed using DENZO [51] and SCALEPACK. Data from the heavy atom derivative were integrated using the program BUDDHA [52] and processed using ROTAVATA and AGROVATA from the CCP4 package [53]. Structure factors from both data sets were calculated using TRUNCATE [53]. Data from the native and derivative were scaled together using SCALEIT [53].

Structure determination and model building

The positions of the heavy atom sites were determined using SHELXS-86 [54]. The positions of the heavy atom sites were refined using MLPHARE [55] and initial SIRAS phases were calculated. The data was then subjected to a round of solvent flattening with histogram matching using DM [56]. A map was calculated which clearly showed the position of the two dimers in the asymmetric unit, and an initial model was built into the initial SIRAS map using the program O [40].

The structure was refined using X-PLOR v3.8.9 [39]. Rounds of positional refinement, followed by simulated annealing and B-factor refinement, were carried out with rebuilding of the structure using O between cycles of refinement. During the initial model building and refinement, omit maps, which excluded 10 residues at a time, were used to check the progress of refinement.

Surface and electrostatic calculations

Surface calculations were performed using the surface option in QUANTA version 4.0. Electrostatic calculations were performed with GRASP version 1.3.

Protein expression and purification

The pR5δV5 plasmid [57], which contains a synthetic gene for HDAg-S, was transformed into *E. coli* BL21(DE3)pLysS cells (Novagen) and purified as described previously [57]. Briefly, 45 ml of frozen cells, corresponding to three 1 l cultures, were thawed and one protease inhibitor tablet (Boehringer Mannheim) was added, as well as RNAse A and DNAse I to a final concentration of 50 μg/ml. Cells lysed by sonication were pelleted at 10,000 × g for 30 min. The lysate was diluted threefold with 50 mM HEPES buffer, pH 7.5, and then applied to a 10 × 1.5 cm Fast SP Sepharose (Pharmacia) column equilibrated with 50 mM HEPES buffer, pH 7.5, and eluted using a salt gradient from 0–1M NaCl in 50 mM HEPES, pH 7.5. The fractions containing rHDAg-S were assayed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and pooled. The sample was then applied to a Superdex S-200 column (Pharmacia) equilibrated with 50 mM Hepes, pH 7.5, 500 mM NaCl and 5% glycerol. The elution of the protein from the column was monitored by UV absorbance at 280 nm.

Mass spectrometry

The rHDAg-S was prepared as described above. The samples for mass spectrometry were prepared as follows: rHDAg-S was dialyzed overnight against water. Cross-linked protein was prepared by the addition of 5 μl of 0.5% glutaraldehyde to 40 μl of rHDAg-S for 5 min, and quenched by the addition of 5 μl of 1 M ammonium acetate. Mass spectrometry was performed in the BCMP Biopolymer facility on a Perseptive Biosystems Voyager-DE mass spectrometer.

Accession numbers

The coordinates have been deposited in the Brookhaven Protein Data Bank (accession number 1A92), and are available from the authors by sending email requests to harmon@dag.med.harvard.edu.

Acknowledgements

We thank members of the Hogle and Ellenberger labs for their help and encouragement, TE Benson and MW Wien for their help in collecting the synchrotron data, Bob Sweet for his assistance at Brookhaven, Russ Henry for peptide synthesis and purification, and Charles Dahl for help with the mass spectrometry. This work was supported by grants from the National Institutes of Health (AI32480 to JMH and R01-HL37974 to SML) and the US Public Health Service (GM 42031 to BWE) and the Giovanni Armenise Harvard Center for Structural Biology.

References

1. Hoofnagle, J.H. (1989). Type D (delta) hepatitis. *JAMA* **261**, 1321-1325.
2. Rizzetto, M., Hoyer, B., Canese, M.G., Shih, J.W., Purcell, R.H. & Gerin, J.L. (1980). delta Agent: association of delta antigen with hepatitis B surface antigen x and RNA in serum of delta-infected chimpanzees. *Proc. Natl Acad. Sci. USA*. **77**, 6124-6128.
3. Kuo, M.Y.P., Chao, M. & Taylor, J. (1989). Initiation of replication of the human hepatitis delta virus genome from cloned DNA: role of delta antigen. *J. Virol.* **63**, 1945-1950.
4. Wang, K.S., *et al.,*& Houghton, M. (1986). Structure, sequence and expression of the hepatitis delta viral genome. *Nature* **323**, 508-513.
5. Kos, A., Dijkema, R., Arnberg, A.C., van der Meide, P. & Schellekens, H. (1986). The hepatitis delta virus possesses a circular RNA. *Nature* **323**, 558-560.
6. Taylor, J. (1990). Hepatitis delta virus: *cis* and *trans* functions needed for replication. *Cell* **61**, 371-373.
7. Wu, H., Lin, Y., Lin, F., Makino, S., Chang, M. & Lai, M.M.C. (1989). Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proc. Natl Acad. Sci. USA.* **86**, 1831-1835.
8. Wu, H.N. & Lai, M.M.C. (1989). Reversible cleavage and ligation of hepatitis delta virus RNA. *Science* **243**, 652-654.
9. Sharmeen, L., Kuo, M.Y., Dinter-Gottlieb, G. & Taylor, J. (1988). Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *J. Virol.* **62**, 2674-2679.
10. Kuo, M., Sharmeen, L., Dinter-Gottlieb, G. & Taylor, J. (1988). Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *J. Virol.* **62**, 4439-4444.
11. Sharmeen, L., Kuo, M. & Taylor, J. (1989). Self-ligating RNA sequences on the antigenome of human hepatitis delta virus. *J. Virol.* **63**, 1428-1430.
12. Makino, S., *et al.*, & Lai, M.M.C. (1987). Molecular cloning and sequencing of a human hepatitis delta (delta) virus RNA. *Nature* **329**, 343-346.
13. Chao, M., Hsieh, S.Y. & Taylor, J. (1990). Role of two forms of hepatitis delta virus antigen: evidence for a mechanism of self-limiting genome replication. *J. Virol.* **64**, 5066-5069.
14. Glenn, J.S. & White, J.M. (1991). *trans*-Dominant inhibition of human hepatitis delta virus genome replication. *J. Virol.* **65**, 2357-2361.
15. Chang, F.L., Chen, P.J., Tu, S.J., Wang, C.J. & Chen, D.S. (1991). The large form of hepatitis delta antigen is crucial for assembly of hepatitis delta virus. *Proc. Natl Acad. Sci. USA.* **88**, 8490-8494.
16. Wang, C.J., Chen, P.J., Wu, J.C., Patel, D. & Chen, D.S. (1991). Small-form hepatitis B surface antigen is sufficient to help in the assembly of hepatitis delta virus-like particles. *J. Virol.* **65**, 6630-6636.
17. Ryu, W.S., Bayer, M. & Taylor, J. (1992). Assembly of hepatitis delta virus particles. *J. Virol.* **66**, 2310-2315.
18. Chang, M.F., Chen, C.J. & Chang, S.C. (1994). Mutational analysis of delta antigen: effect on assembly and replication of hepatitis delta virus. *J. Virol.* **68**, 646-653.
19. Chang, M.F., *et al.,* & Lai, M.M.C. (1988). Human hepatitis delta antigen is a nuclear phosphoprotein with RNA-binding activity. *J. Virol.* **62**, 2403-2410.
20. Lin, J.H., Chang, M.F., Baker, S.C., Govindarajan, S. & Lai, M.M.C. (1990). Characterization of hepatitis delta antigen: specific binding to hepatitis delta virus RNA. *J. Virol.* **64**, 4051-4058.
21. Chao, M., Hsieh, S.Y. & Taylor, J. (1991). The antigen of hepatitis delta virus: examination of *in vitro* RNA-binding specificity. *J. Virol.* **65**, 4057-4062.
22. Xia, Y.P. & Lai, M.M.C. (1992). Oligomerization of hepatitis delta antigen is required for both the *trans*-activating and *trans*-dominant inhibitory activities of the delta antigen. *J. Virol.* **66**, 6641-6648.
23. Wang, J.G. & Lemon, S.M. (1993). Hepatitis delta virus antigen forms dimers and multimeric complexes *in vivo*. *J. Virol.* **67**, 446-454.
24. Chang, M.F., Sun, C.Y., Chen, C.J. & Chang, S.C. (1993). Functional motifs of delta antigen essential for RNA binding and replication of hepatitis delta virus. *J. Virol.* **67**, 2529-2536.

25. Xia, Y.P., Yeh, C.T., Ou, J.H. & Lai, M.M.C. (1992). Characterization of nuclear targeting signal of hepatitis delta antigen: nuclear transport as a protein complex. *J. Virol.* **66**, 914-921.
26. Lee, C.Z., Lin, J.H., Chao, M., McKnight, K. & Lai, M.M.C. (1993). RNA-binding activity of hepatitis delta antigen involves two arginine-rich motifs and is required for hepatitis delta virus RNA replication. *J. Virol.* **67**, 2221-2227.
27. Lazinski, D.W. & Taylor, J.M. (1993). Relating structure to function in the hepatitis delta virus antigen. *J. Virol.* **67**, 2672-2680.
28. Lee, C.Z., Chen, P.J., Lai, M.M.C. & Chen, D.S. (1994). Isoprenylation of large hepatitis delta antigen is necessary but not sufficient for hepatitis delta virus assembly. *Virology* **199**, 169-175.
29. Glenn, J.S., Watson, J.A., Havel, C.M. & White, J.M. (1992). Identification of a prenylation site in delta virus large antigen. *Science* **256**, 1331-1333.
30. Hwang, S.B. & Lai, M.M.C. (1993). Isoprenylation mediates direct protein–protein interactions between hepatitis large delta antigen and hepatitis B virus surface antigen. *J. Virol.* **67**, 7659-7662.
31. de Bruin, W., Leenders, W., Kos, T. & Yap, S.H. (1994). *In vitro* binding properties of the hepatitis delta antigens to the hepatitis B virus envelope proteins: potential significance for the formation of delta particles. *Virus Res.* **31**, 27-37.
32. Hwang, S.B. & Lai, M.M.C. (1993). A unique conformation at the carboxyl terminus of the small hepatitis delta antigen revealed by a specific monoclonal antibody. *Virology* **193**, 924-931.
33. Hwang, S.B. & Lai, M.M.C. (1994). Isoprenylation masks a conformational epitope and enhances *trans*-dominant inhibitory function of the large hepatitis delta antigen. *J. Virol.* **68**, 2958-2964.
34. Chen, P.J., Chang, F.L., Wang, C.J., Lin, C.J., Sung, S.Y. & Chen, D.S. (1992). Functional study of hepatitis delta virus large antigen in packaging and replication inhibition: role of the amino-terminal leucine zipper. *J. Virol.* **66**, 2853-2859.
35. MacNaughton, T.B., Gowans, E.J., McNamara, S.P. & Burrell, C.J. (1991). Hepatitis delta antigen is necessary for access of hepatitis delta virus RNA to the cell transcriptional machinery but is not part of the transcriptional complex. *Virology* **184**, 387-390.
36. Fu, T.-B. & Taylor, J. (1993). The RNAs of hepatitis delta virus are copied by RNA polymerase II in nuclear homogenates. *J. Virol.* **67**, 6965-6972.
37. Rozzelle, J.E., Jr., Wang, J.G., Wagner, D.S., Erickson, B.W. & Lemon, S.M. (1995). Self-association of a synthetic peptide from the N terminus of the hepatitis delta virus protein into an immunoreactive alpha-helical multimer. *Proc. Natl Acad. Sci. USA.* **92**, 382-386.
38. Wang, J.-G., Jansen, R.W., Brown, E.A. & Lemon, S.M. (1990). Immunogenic domains of hepatitis delta virus antigen: peptide mapping of epitopes recognized by human and woodchuck antibodies. *J. Virol.* **64**, 1108-1116.
39. Brünger, A.T. (1992). *A System for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT.
40. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1990). Improved methods of building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110-119.
41. Talluru, S.R., Krishna, X.-P.K., Gary, S., Burgers, P.M. & Kuryian, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**, 1233-1243.
42. Chang, M.F., Chang, S.C., Chang, C.I., Wu, K. & Kang, H.Y. (1992). Nuclear localization signals, but not putative leucine zipper motifs, are essential for nuclear transport of hepatitis delta antigen. *J. Virol.* **66**, 6019-6027.
43. Poisson, F., *et al.*, & Goudeau, A. (1993). Characterization of RNA-binding domains of hepatitis delta antigen. *J. Gen. Virol.* **74**, 2473-2478.
44. Poisson, F., Roingeard, P. & Goudeau, A. (1995). Direct investigation of protein RNA-binding domains using digoxigenin-labelled RNAs and synthetic peptides: application to the hepatitis delta antigen. *J. Virol. Methods* **55**, 381-389.
45. Lilley, D.M.J. (1995). Kinking of DNA and RNA by base bulges. *Proc. Natl Acad. Sci. USA.* **92**, 7140-7142.
46. Brazas, R. & Ganem, D. (1996). A cellular homolog of hepatitis delta antigen – implications for viral replication and evolution. *Science* **274**, 90-94.
47. Soisson, S.M., MacDougall-Shackleton, B., Schleif, R. & Wolberger, C. (1997). Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421-425.
48. Bussiere, D.E., Bastia, D. & White, S.W. (1995). Crystal structure of the replication terminator protein from *B. Subtilis* at 2.6 Angstrom. *Cell* **80**, 651-660.
49. Woolfson, D.N. & Alber, T. (1995). Predicting oligomerization states of coiled coils. *Protein Sci.* **4**, 1596-1607.
50. Wolf, E., Kim, P.S. & Berger, B. (1997). MULTICOIL - A program for predicting two and three stranded coiled coils. *Protein Sci.* **6**, 1179-1189.
51. Otwinowski, Z. (1993). *Data Collection and Processing*. pp. 56-62, SERC Daresbury Laboratory, Warrington, UK.
52. Blum, M., Metcalf, P., Harrison, S.C. & Wiley, D.M. (1987). A system for collection and on-line integration of X-ray diffraction data from a multiwire detector. *J. Appl. Cryst.* **20**, 235-242.
53. CPP4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760-763.
54. Sheldrick, G.M. (1990). Phase annealing in SHELX-90: direct methods for larger structures. *Acta Cryst. A* **46**, 467-473.
55. Otwinowski, Z. (1991). Isomorphous replacement and anomalous scattering. In *Proceedings of the CCP4 Study Weekend*. pp. 80-86. SERC Daresbury Laboratory, Warrington, UK.
56. Zhang, K.Y.J. & Main, P. (1990). Histogram matching as a new density modification technique for phase refinement and extension of protein molecules. *Acta Cryst. A* **46**, 41-46.
57. Dingle, K., Bichko, V., Zuccola, H., Hogle, J. & Taylor, J. (1998). Initiation of hepatitis delta virus genome replication. *J. Virol.* **72**, 4783–4788.
58. Carson, M. & Bugg, C.E. (1986). Algorithm for ribbon models of proteins. *J. Mol. Graph.* **4**, 121-122.
59. Kraulis, P.J. (1991). Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.
60. Nicholls, A. (1992). *GRASP: Graphical Representation and Analysis of Surface Properties.* Columbia University, New York, NY.
61. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0, a program for photorealistic molecular graphics. *Acta Cryst. D* **50**, 869-873.