The EMBO Journal Vol.16 No.19 pp.5819–5826, 1997

The three-dimensional solution structure of the matrix protein from the type D retrovirus, the Mason–Pfizer monkey virus, and implications for the morphology of retroviral assembly

Maria R.Conte[1], Michaela Klikova[2], Eric Hunter[3], Tomas Ruml[2] and Stephen Matthews[1,4]

[1]Department of Biochemistry, Imperial College of Science, Technology and Medicine, University of London, London SW7 2AY, UK, [2]Department of Biochemistry and Microbiology, Institute of Chemical Technology, Prague, Czech Republic and [3]Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[4]Corresponding author
e-mail: s.j.matthews@ic.ac.uk

The Mason–Pfizer monkey virus (M-PMV) is the prototype of the type D retroviruses. In type B and D retroviruses, the Gag protein pre-assembles before association with the membrane, whereas in type C retroviruses (lentiviruses, BLV/HTLV group) Gag is targeted efficiently to the plasma membrane, where the particle formation occurs. The N-terminal domain of Gag, the matrix protein (MA), plays a critical role in determining this morphogenic difference. We have determined the three-dimensional solution structure of the M-PMV MA by heteronuclear nuclear magnetic resonance. The protein contains four α-helices that are structurally similar to the known type C MA structures. This similarity implies possible common assembly units and membrane-binding mechanisms for type C and B/D retroviruses. In addition to this, the interpretation of mutagenesis data has enabled us to identify, for the first time, the structural basis of a putative intracellular targeting motif.

Keywords: Mason–Pfizer monkey virus/matrix protein/NMR/retroviral assembly/solution structure

Introduction

Gag proteins, the major structural components of retroviruses, play a central role in viral replication at both the early and late stages of the virus life cycle (Wills and Craven, 1991; Hunter, 1994). The *gag* gene of mammalian retroviruses encodes a precursor protein which is responsible for the assembly and release of the virion particle from the infected cells. Shortly after budding or in the late stage of the budding, the polyprotein precursor is proteolytically processed by the viral protease into the principle mature Gag-derived proteins: matrix (MA), capsid and nucleocapsid. The N-terminal matrix protein of the *gag* gene product has been found associated with the virus envelope glycoproteins in most mammalian retroviruses (Gerderblom *et al.*, 1989; Arnold and Arnold, 1991; Nermut *et al.*, 1994; Spearman *et al.*, 1994). Several lines of evidence suggest that this protein is involved in virus particle assembly, transport and budding, even though the processes of retroviral assembly and maturation and the exact roles of the Gag proteins in these events are not fully understood. In type C retroviruses [lentiviruses, bovine leukaemia virus (BLV)/human T-lymphotrophic virus (HTLV) group], the MAs are targeted efficiently to the plasma membrane, where the particle formation occurs. On the contrary, in type B and D retroviruses, the Gag protein pre-assembles before association with the membrane (Gerderblom, 1991; Wills and Craven, 1991; Hunter, 1994).

The Mason–Pfizer monkey virus (M-PMV) is the prototype of the type D retroviruses. This virus is one of the best characterized simian retroviruses (SRV) (Van Der Kuyl *et al.*, 1997) which induces a fatal immmunodeficiency syndrome in rhesus macaques, initially called SAIDS for simian AIDS (Marx *et al.*, 1984; Stromberg *et al.*, 1984). Nevertheless, the SRVs are unrelated to SIV (simian immmunodeficiency virus), which currently is recognized as the simian counterpart of the human immmunodeficiency virus.

Although M-PMV normally assembles its interior protein shell, the capsid, in the cytoplasm, a single amino acid substitution in the M-PMV MA resulted in virus particle assembly at the plasma membrane rather than within the cytoplasm, in a similar way to type C retroviruses (Rhee and Hunter, 1990). The mutant, which has one amino acid substitution at position 55 (arginine to tryptophan), allowed the identification of a dominant signal in the Gag of M-PMV responsible for the intracytoplasmic capsid assembly in type D retroviruses. This signal appears to prevent immediate transport of individual Gag proteins to the plasma membrane or targets them to an intracellular assembly site. This, together with other studies (Yu *et al.*, 1992; Fäcke *et al.*, 1993;Yuan *et al.*, 1993; Freed *et al.*, 1994), implies a key role for the MA in the morphogenesis of retroviruses.

The sequence homology between MAs of different mammalian retroviruses is very low, but most of them show the common feature of a myristate attached at the N-terminal glycine (after the removal of the terminal methionine). The high resolution three-dimensional structures of several MAs of type C retroviruses (HIV, HTLV, BLV and SIV) have been solved by both crystallography and NMR (Massiah *et al.*, 1994; Matthews *et al.*, 1994, 1995, 1996; Rao *et al.*, 1995; Christensen *et al.*, 1996; Hill *et al.*, 1996). Despite the low sequence homology, the structures of HIV, HTLV, BLV and SIV MAs were found to be very similar, which could imply a common assembly design principle for the MA of type C retroviruses (Matthews *et al.*, 1996).

Here we report the high resolution 3D structure of the MA of the M-PMV, which we believe is the first structural insight into an MA from a type D retrovirus. The structure of M-PMV MA shows four α-helices closely packed

together which are topologically related to HIV, SIV, BLV and HTLV MAs. In addition, M-PMV MA shows the characteristic localized patch of basic residues, commonly seen in type C retroviruses, which has been suggested to be responsible for the efficient targeting to and association with the membrane (Yuan *et al.*, 1993; Zhou *et al.*, 1994; Matthews *at al.*, 1996).

Mutagenesis and specific in-frame deletion analyses within the MA of M-PMV and HIV have provided basic information about the role of the MA in Gag stability and in viral capsid assembly for both types of retroviruses. As far as M-PMV is concerned, several key modifications have been identified, some of them affecting the capsid assembly, others playing crucial roles in capsid transport and membrane association (Rhee and Hunter, 1990, 1991). We have been able to provide structural insights that allow interpretation of these data and we suggest plausible models for the different morphogenesis between type C and type B/D retroviruses.

Results and discussion

The M-PMV MA was expressed in *Escherichia coli*. The NMR sequence-specific backbone ^{1}H assignments were obtained using 3D ^{15}N–^{1}H nuclear Overhauser (NOESY) heteronuclear multiple quantum coherence (HMQC) and Hartmann–Hahn (HOHAHA) HMQC spectroscopy (Marion *et al.*, 1989; Driscoll *et al.*, 1990). The secondary structure elements and ^{1}H side-chain assignments subsequently were determined. Figure 1 shows representative strips taken from 3D NOESY HMQC for the region between Thr21 and Leu31. The distance restraints used in the structure calculations were obtained from the NOESY HMQC experiment and 2D homonuclear NOESY spectra. The structures were calculated within the program XPLOR using a dynamic simulated annealing protocol (Nilges *et al.*, 1988; Brunger, 1993) on the basis of 1069 nuclear Overhauser effect (NOE) distance restraints and 45 H bond distance restraints.

Solution structure of M-PMV MA

A final family of 15 refined structures has been defined (Figure 2) for M-PMV MA. The overall root-mean-square deviation (r.m.s.d.) between the family and the mean coordinate position is 0.79 ± 0.1 Å for the backbone atoms and 1.31 ± 0.1 Å for all the heavy atoms in regions of secondary structures (0.97 ± 0.1 Å for the backbone atoms and 1.56 ± 0.1 Å for all the heavy atoms for all residues between 6 and 94). The structural calculation statistics are shown in Table I. M-PMV MA shows four principal α-helices closely packed together (Figure 2), joined by short loops or regions of extended structure. Helix A runs from residue His6 to Gly23, making very close hydrophobic contacts with helix B, spanning residues 29–42, which finishes in a well-defined 3-10 helix from residue Pro42 to Phe45. Helix B is joined to helix C by a semi-structured loop, Gln47–Ile51, forming a tight turn centred on Gly49, with some evidence of β-sheet-like NOEs across the chains. Helix C, running from Ile53 to Thr69, ends in a short loop showing NOE evidence for a 3-10 helix between residues 72 and 75 that is not well defined within the family of structures. The C-terminal helix, helix D, runs from Val77 to Glu94 and forms close

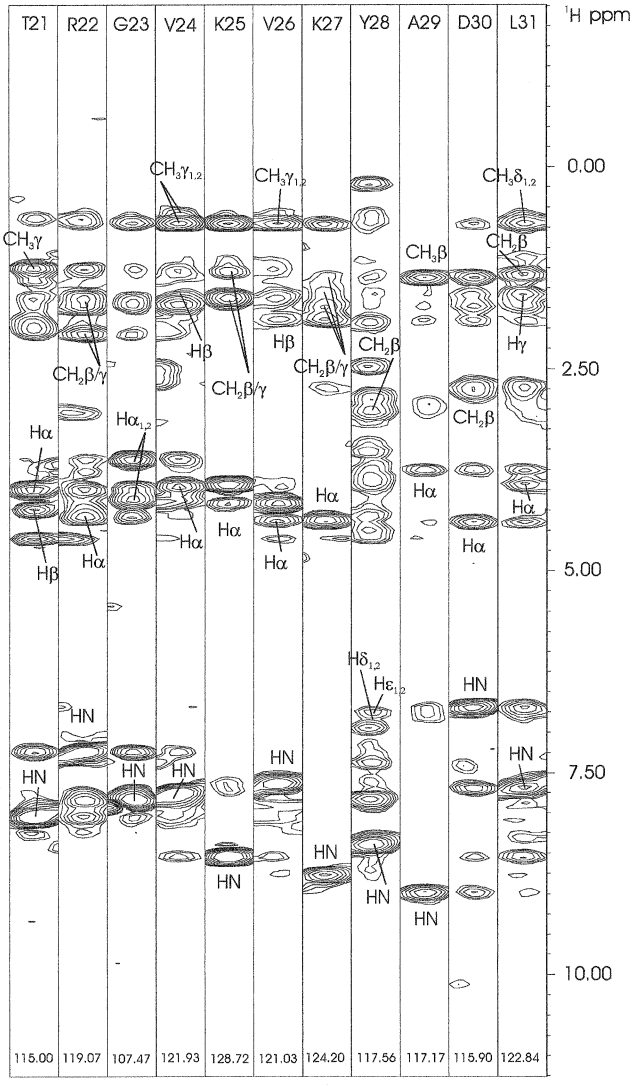

Fig. 1. Strips taken from the 3D ^{15}N-^{1}H nuclear Overhauser (NOESY) heteronuclear multiple quantum coherence (HMQC) experiment, for the region spanning residues 21–31. The assigned protons are labelled. The ^{15}N chemical shift is indicated for each amide at the bottom of each strip.

contacts with helix C. Helices A and B lie approximately perpendicular to helices C and D, with numerous interhelical contacts defining an extensive hydrophobic core. The major contacts involve Tyr11, Leu15 and Leu19 from helix A, Tyr28, Leu31, Leu32, Phe34, Phe35, Phe37 and Val38 from helix B, Ile53, Trp56, Val59, Phe63, Tyr66 and Tyr67 from helix C, and Ala79, Phe80, Tyr82, Trp83, Leu85, Ile86, Leu89 and Ile90 from helix D.

In M-PMV MA, the four proline residues at positions 43, 46, 72 and 76 provide the structural basis for semi-structured loops and the initiation of two 3-10 helices spanning the regions 42–45 and 72–75 respectively.

The structure of M-PMV MA shows two distinct positively charged regions located on opposite sides of the molecule. The first region is analogous to the N-terminal basic region seen in the structures of HIV, SIV and BLV MA. More specifically, it contains side chains of residues from helices A and B and from the loop C–D, namely Arg10, Lys16, Lys20, Arg22, Lys25, Lys27, Lys33, Lys39 and Lys74. Side chains of residues Lys54,

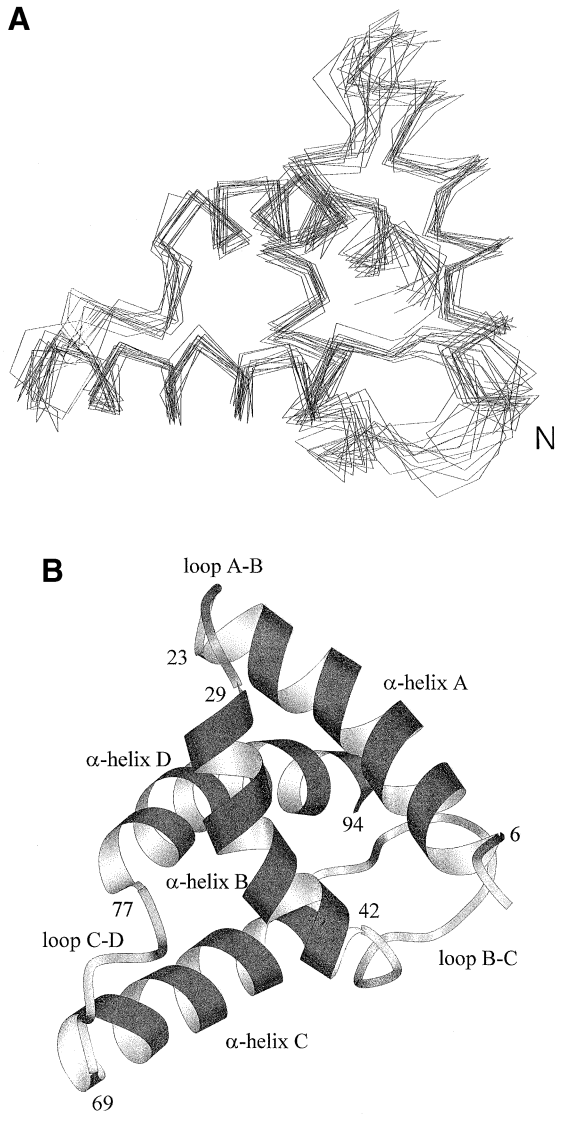

Fig. 2. (**A**) Superimposition of the 15 refined M-PMV MA structures. The Cα trace for residues 6–94 in BLV MA. (**B**) A MOLSCRIPT (Kraulis, 1991) diagram representing the backbone trace. The start and stop residues of the helices are numbered.

Arg55, Arg57, Arg58, Lys87, Lys92 and Lys93 of helices C and D contribute to the second positively charged surface. This particular feature has not been observed for type C retroviral MAs.

Structural similarity with type C MAs and implications for assembly

An automatic search of the Brookhaven protein structure database with the program DALI was used to investigate the structural similarity of M-PMV MA to other proteins (Holm and Sander, 1993). The fold of M-PMV MA has been found to be extremely similar to the MAs from the type C retroviruses: HIV, SIV, BLV and HTLV MAs (Figure 3). A summary of the superimposition results is shown in Table II. The first four α-helices of HIV and SIV MA (Matthews *et al.*, 1994, 1995; Rao *et al.* 1995) display a similar topology to that of the M-PMV MA α-helices. In addition, the 3-10 helix at 49–52 in HIV MA aligns well with that the 42–45 3-10 helix of M-PMV MA. The superimposition of all the aligned regions of

Table I. Results from structure calculations for M-PMV MA

Statistic		<SA>[a]
Restraints (°)		
Intra-residue	248	0.028 ± 0.01
Medium and short range (i to $1<i<5$)	612	0.032 ± 0.01
Long range (i to $i>4$)	209	0.056 ± 0.003
Hydrogen bonds	45	0.016 ± 0.004
R.m.s.d. from restraints and idealized geometry		
NOE distances (Å)		0.036 ± 0.004
Bonds (Å)		0.0037 ± 0.0002
Angles (°)		0.68 ± 0.02
Improper angles (°)		0.50 ± 0.05
Final experimental energy terms (kJ/mol)		
F_{NOE}		78.9 ± 8.2
F_{angle}		203.5 ± 12.0
$F_{improper}$		33.4 ± 4.2
F_{bond}		22.4 ± 1.8
R.m.s.d. from coordinate positions		
Backbone atom for helices		0.79 ± 0.1[b]
All atoms for helices		1.31 ± 0.1
Backbone atoms between residues 6 and 94		0.97 ± 0.1
All atoms between residues 6 and 94		1.56 ± 0.1

[a]The average r.m.s.d. for the final 15 structures.
[b]The average r.m.s.d. from the average structure.

M-PMV MA and HIV MA gave an r.m.s.d. of 3.9 Å over 73 Cα atoms. The β-sheet region of HIV and SIV MA, which contains the nuclear localization signal, is not present in M-PMV MA. The general helical topology of M-PMV MA is more reminiscent of that of BLV and HTLV MAs (Christensen *et al.*, 1996; Matthews *et al.*, 1996). The major significant differences between these two structures and that of M-PMV MA are the orientation of the loop B–C and the length of the helices C and D.

Based upon the similar helical topology and loop conformation between M-PMV and HIV MAs, we propose that M-PMV MA can form a trimer as a fundamental intermediate of Gag assembly, as observed in the crystal structures of HIV and SIV MAs. The putative M-PMV MA trimer, built by superimposition of M-PMV MA monomers on the HIV/SIV trimer, shows the interface loops in the regions around residues 44 and 69 (Figure 4A and B). As with type C MAs, ^{1}H and ^{15}N linewidths indicate that M-PMV MA is predominantly monomeric in solution. However, there is some experimental evidence to suggest that M-PMV MA does weakly associate with itself. Significant ^{1}H and ^{15}N chemical shift changes of amides (data not shown) were observed as a function of protein concentration. Moreover, these amides were located predominantly in the trimer interface regions.

The HIV, SIV and BLV trimer models present all the N-terminal residues implicated in membrane targeting and binding on the upper surface. This provides a membrane-interacting site that contains the *N*-myristate group for burial within the membrane and the basic side chains for association with the acid phospholipid head groups. The membrane-interacting site of the M-PMV MA trimer appears to be very similar to that of the type C retroviruses, showing the N-terminal basic residues on the upper surface and the N-terminus appropriately positioned for myristate

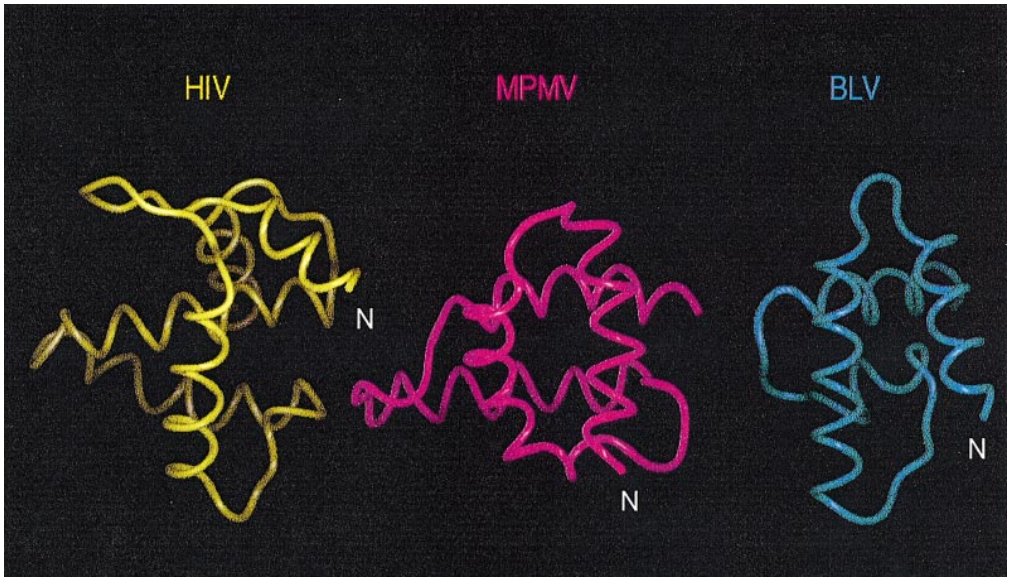

Fig. 3. The structural similarity between the MAs from HIV, BLV and M-PMV. Protein worms showing a superimposition of HIV, BLV and M-PMV MA. For clarity, the molecules are separated.

burial. It is probable that, on the basis of the similarities between M-PMV and HIV MA structures, the process of Gag transport and binding to the plasma membrane in type C and D retroviruses is highly related, and in both cases the common feature of a bipartite membrane association signal appears to be an absolute requirement.

The other basic region observed in M-PMV MA, which is not a feature of type C retroviral MAs, is located on the lower surface of the trimer (Figure 4). The functional significance of this is unclear, but it may possibly interact with the negative charges of the highly phosphorylated second Gag protein, pp24, (Bradac *et al.*, 1985), or the acidic N-terminus of p12, a Gag domain involved in intracytoplasmic capsid assembly (Sommerfelt *et al.*, 1992), and absent in HIV.

Table II. Summary of superimposition statistics for M-PMV MA with other retroviral matrix proteins

	BLV[a]	HTLV[b]	HIV[c]
R.m.s.d. (Å)	3.3	3.5	3.9
No. of Cα atoms	64	68	73
Identity/%	9	4	5

[a]Residues 6–13, 15–20, 25–43, 44–47, 59–62, 63–66, 75–93 in M-PMV MA with 16–23, 24–29, 30–48, 52–55, 60–63, 65–68, 69–87 in BLV MA.
[b]Residues 5–10, 11–27, 30–43, 48–52, 54–59, 61–64, 77–88, 90–93 in M-PMV MA with 16–21, 23–39, 41–54, 62–66, 67–72, 73–76, 78–89, 90–93 in HTLV MA.
[c]Residues 11–14, 15–23, 27–40, 41–45, 51–64, 66–71 73–93 in M-PMV MA with 7–10, 12–20, 30–43, 45–49, 52–65, 66–71, 72–92 in HIV MA.

Interpretation of mutagenesis data and implications for membrane targeting

Several amino acid substitutions with the M-PMV MA have been found to affect assembly, transport and membrane association of capsids (Rhee and Hunter, 1991). The M-PMV MA appears, therefore, to play key roles in each of these sequential events in the late stages of the virus life cycle, and mutants defective in these steps have been characterized. In an attempt to rationalize the phenotypic effects of these mutations on M-PMV assembly, individual mutations have been analysed with respect to their locations in the M-PMV MA structure. Table III summarizes the published mutations (Rhee and Hunter, 1991) and their consequences on the final stages on the virus cycle. Figure 5A and B shows the location of some of the mutants within the proposed M-PMV MA trimer model. Mutants T21I, T41I and R57C have no effect on assembly. These side chains are highly exposed and thus, in general, are more tolerant to alteration. Mutations of prolines 43 and 72 cause a reduction in Gag stability and particle assembly. These residues are located in exposed loops within the monomer that form part of the interface in the proposed trimer model (Figure 5A). It is, therefore, highly likely that these modifications would substantially alter the conformation of the loops. This, in turn, would not only affect the stability of the molecule but also disrupt formation of the proposed trimer intermediate, resulting in disturbed particle assembly.

One of these single point mutations, Arg55→Trp, provides direct evidence for the key role of MA in the morphogenesis of retroviruses, since it redirects the assembly of capsids from an intracytoplasmic site to the plasma membrane in a manner similar to that of type C retroviruses (Rhee and Hunter, 1990). The fact that a single point mutation can significantly alter morphogenesis in this way suggests that the difference between type C and type D retrovirus capsid assembly is very subtle. This is supported by the high degree of structural similarly found for the M-PMV MA and the MAs from viruses exhibiting type C morphogenesis. The structure of the M-PMV MA shows that the arginine residue at position 55 is exposed and located towards the end of helix B in an 18 amino acid region, spanning residues 43–60 (Figure 5C), that is highly conserved only in type D (M-PMV) and type B (mouse mammary tumor virus; MMTV) retroviruses (Rhee and

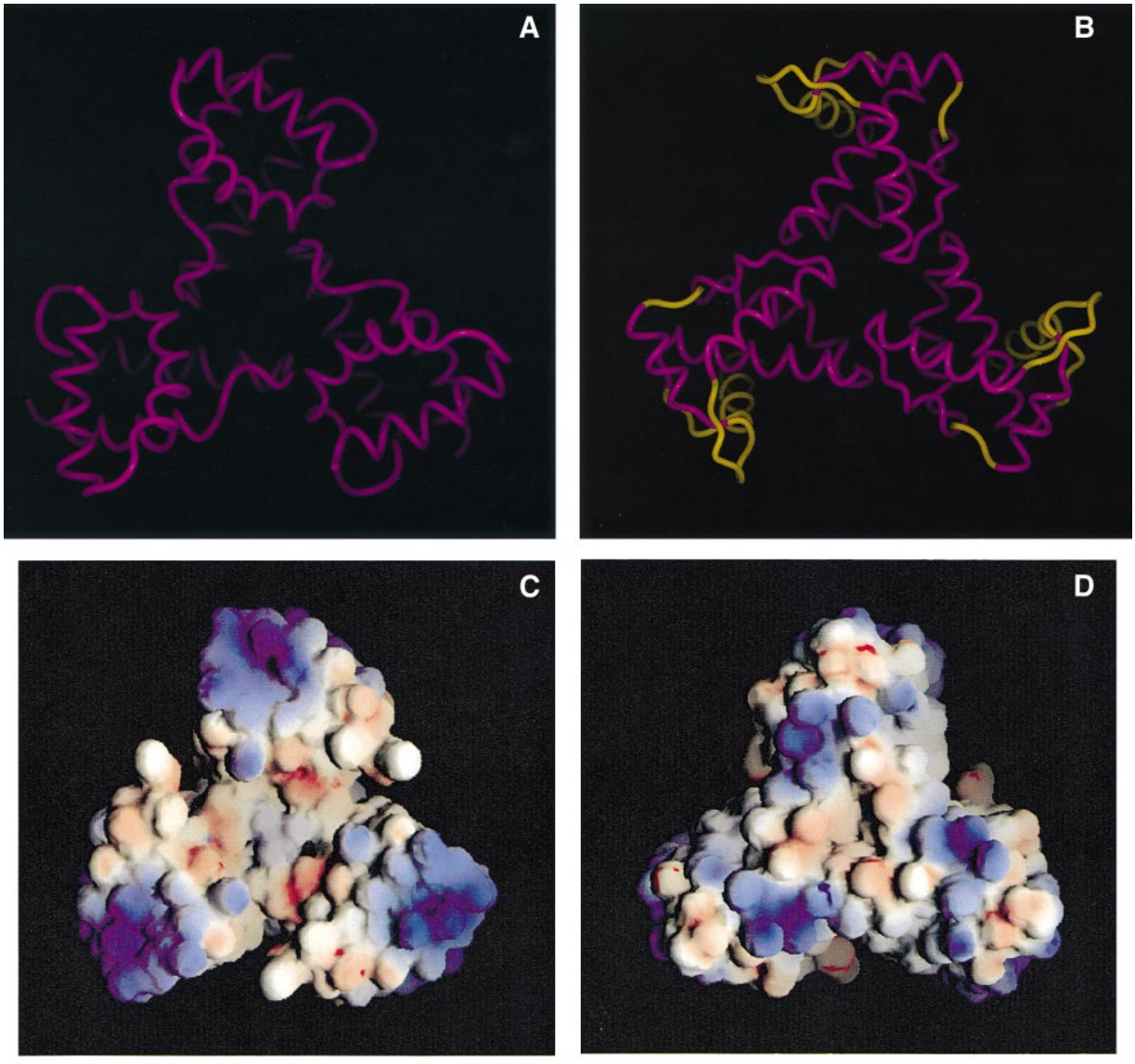

Fig. 4. Representation of the trimer model for M-PMV MA. (**A**) Protein worm showing the relative orientation of the monomers of M-PMV MA within the trimer model. (**B**) Protein worm showing the relative orientation of the monomers of HIV MA within the trimer model. The regions structurally similar to M-PMV MA are shown in magenta and the rest in yellow. (**C**) A GRASP representation (Nicholls *et al.*, 1991) of the electrostatic properties of the M-PMV MA trimer shown from the top. The calculated dipole is perpendicular to the membrane. Blue colour indicates positive charge and red negative. (**D**) A GRASP representation (Nicholls *et al.*, 1991) of the electrostatic properties of the M-PMV MA trimer shown from the bottom. The calculated dipole is perpendicular to the membrane. Blue colour indicates positive charge and red negative.

Table III. Summary of mutagenesis data for M-PMV MA

Mutation	Side chain location		Mutant behaviour				
	Monomer	Trimer	WT	Stability and assembly	Transport deficient	Membrane binding deficient	Type C
T(21)→I	exposed hydrophilic,	exposed, hydrophilic	■				
T(41)→I	exposed, hydrophilic	exposed, hydrophilic	■				
R(57)→C	exposed, charged	exposed, charged	■				
P(43)→L	exposed, hydrophobic	trimer interface		■			
P(72)→S	exposed, hydrophobic	trimer interface		■			
del-Myr	unknown	unknown			■		
A(18)→V	buried, hydrophobic	buried, hydrophobic			■		
A(79)→V	buried, hydrophobic	buried, hydrophobic			■		
T(69)→I	exposed, hydrophilic	exposed, hydrophilic			■		
T(78)→I	exposed, hydrophilic	exposed, hydrophilic				■	
R(55)→W	exposed, hydrophilic	exposed, hydrophilic					■

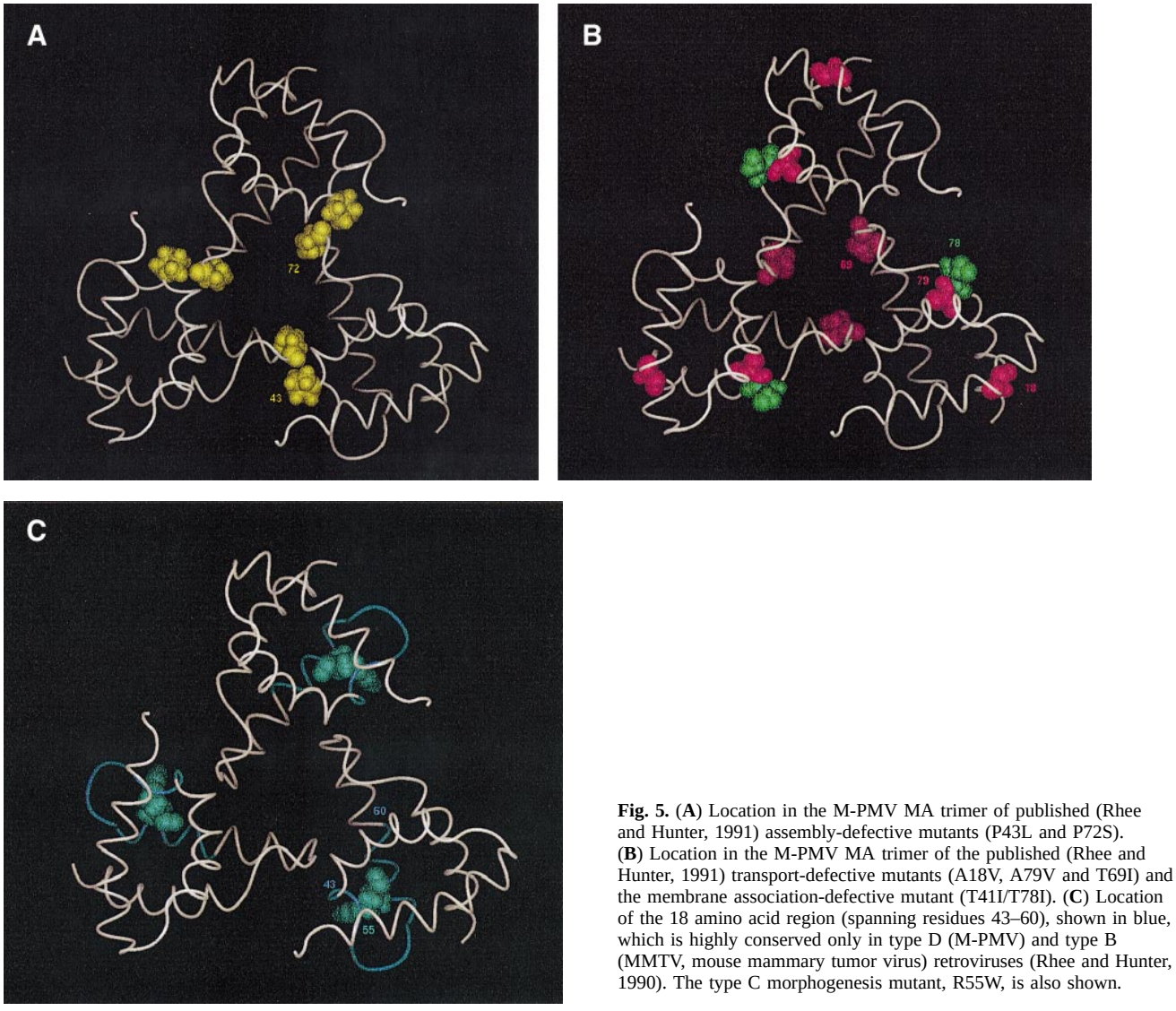

Fig. 5. (**A**) Location in the M-PMV MA trimer of published (Rhee and Hunter, 1991) assembly-defective mutants (P43L and P72S). (**B**) Location in the M-PMV MA trimer of the published (Rhee and Hunter, 1991) transport-defective mutants (A18V, A79V and T69I) and the membrane association-defective mutant (T41I/T78I). (**C**) Location of the 18 amino acid region (spanning residues 43–60), shown in blue, which is highly conserved only in type D (M-PMV) and type B (MMTV, mouse mammary tumor virus) retroviruses (Rhee and Hunter, 1990). The type C morphogenesis mutant, R55W, is also shown.

Hunter, 1990). Since in the MA of murine leukaemia virus (MuLV), the residues equivalent to 50–55 in M-PMV MA are deleted, while flanking homologies are conserved, it is tempting to speculate that this short stretch of amino acids represents a targeting signal, which specifically directs Gag to an intracellular assembly site. Figure 5C shows that this amino acid domain encompasses the structured turn centred at residue 49 and protrudes into the solvent in both monomer and trimer models. This region would, therefore, be ideally positioned for recognition by a specific intracellular transport factor. Moreover, this loop region is shorter and adopts a different conformation in the structures of each of the type C MAs solved to date. It is also interesting to point out that the mutation at position 57 also lies proximal to this loop region. However, mutation of this exposed arginine to the smaller cysteine is more likely to have a negligible effect on the loop conformation. The arginine at position 55 is positioned closer to the loop and its modification to the larger and more hydrophobic side chain of tryptophan will probably influence the conformation of the surrounding region.

One of the most interesting modifications to MA is the deletion of the N-terminal myristate group. In HIV and other type C retroviruses (Freed *et al.*, 1994), this results in defective membrane binding and abrogation of particle assembly. Zhou and Resh (1996) have proposed that, when the myristate is exposed, it enables HIV Gag to bind to the plasma membrane during assembly, and that cleavage into the mature MA promotes the burial of the myristate group, which might disrupt the interaction with the membrane and initiate localization to the nucleus. In HIV, the myristate group, together with a series of N-terminal basic amino acids, has been shown to be crucial for plasma membrane targeting (Zhou *et al.*, 1994; Gallay *et al.*, 1995). In M-PMV, removal of the myristate group does not prevent capsid assembly but blocks the intracellular transport of assembled capsids to the plasma membrane and results in their accumulation at distinct sites within the cytoplasm (Rhee *et al.*, 1990). Thus the structure presented here, which represents the non-myristylated form of MA, is, in the context of the intact Gag molecule, a transport-defective form of the protein. Similarly, these data show that while M-PMV MA presents a series of

basic residues on the upper surface of the trimer, in an analogous manner to that observed for HIV MA, in the absence of myristate this alone is insufficient to mediate transport to the plasma membrane. It is likely, therefore, that the myristate moiety of M-PMV Gag modifies the capsid structure in order to play its integral role in the capsid's progression to the membrane. While it is possible that this could reflect the exposure of myristate itself, as suggested by Zhou and Resh (1996), in other viruses where the structure has been determined, the myristate is buried within the capsid (Moscufo and Chow, 1992; Smyth *et al.*, 1995). It will be interesting, therefore, to determine whether the myristate of M-PMV Gag might pack within the hydrophobic core.

An indirect role for myristate in intracellular transport is consistent with the observation that mutations which do not prevent myristate addition to Gag can nevertheless block progression from the site of assembly. These mutants that have altered transport (A18V, A79V and T69I) presumably have a similar effect on myristate accessibility or capsid structure. These mutants are distributed throughout the M-PMV MA (Figure 5). Mutants A18V and A79V, which are severely defective for transport and accumulate large inclusions of pre-assembled capsids (Rhee and Hunter, 1991), both contain mutations that introduce a more bulky hydrophobic residue into the intra-helix hydrophobic core, conceivably affecting myristate accessibility. Mutant T69I, which transports assembled capsids more slowly to the plasma membrane (Rhee and Hunter, 1991), results from the substitution of a hydrophobic residue for a polar threonine close to the potential trimer interface. Since capsids are assembled readily, this change cannot be significantly damaging to the stability or structure of M-PMV MA. However, it is likely to have subtle effects on the global capsid structure that could be enough to disrupt the mechanism of transport, whether the myristate group accessibility is altered or a conformational change prevents recognition.

One additional double mutant of M-PMV, T41I/T78I, has a unique phenotype. This mutant assembles capsids and transports them to the plasma membrane with wild-type kinetics, but is defective in initiating the process of budding through the membrane. Thus large numbers of capsids can be found lining the inner side of the plasma membrane (Rhee and Hunter, 1991). These amino acid substitutions at the C-terminus of helix B and the N-terminus of helix D introduce hydrophobic residues that point into the intra-helix core. If the myristate moiety is sequestered within this core, such changes might stabilize the hydrophobic interactions, preventing exposure of the myristate to the lipid bilayer and, according to the model of Zhou and Resh (1996), reduce stable membrane binding. Type D and type B capsids differ from other retroviruses, in that following transport to the plasma membrane (and possible charge interactions with the polar head groups), they must wrap a pre-formed capsid with the lipid bilayer during the process of budding. It is tempting to speculate that myristate–lipid interactions might mediate this process and that the T41I/T78I mutant is defective in this step.

In conclusion, we propose that the major principles of type C and type D MA design are based around a few common requirements: a trimeric assembly unit that presents a positively charged surface, and a myristate group for transport and anchorage to the membrane. In addition to this, type B/D MAs present a putative intracellular targeting motif for recognition.

Materials and methods

MPMV MA expression and purification

The vector pG10MA was constructed by deletion of a 1.6 kbp *Eco*57–*Sph*I fragment from the *gag* region of the expression vector pG10GAG which was described previously (Klikova *et al.*, 1995). The resulting vector contained the MA sequence followed by the flanking 15 amino acids from Gag.

For production of MA, the appropriate plasmid was introduced into competent cells of *E.coli* BL21(DE3). The expression of M-PMV proteins in *E.coli* was induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.4 mM. The cells were harvested 4 h after induction. The cells were lysed with lysosyme in Tris–HCl (pH 8.0, 50 mM) buffer containing 1 mM EDTA and 100 mM NaCl. The cell lysate was incubated for 30 min at 4°C in the presence of sodium deoxycholate (0.1%). DNase I (1 mg/ml) and RNase A (2 mg/ml) were added and the lysate was incubated for an additional 30 min at room temperature. In order to cleave off the flanking 15 amino acid sequence, protein was incubated for 1 h with recombinant M-PMV protease (Pichova *et al.*, 1997). The debris was pelleted by centrifugation. MA was precipitated with 75% ammonium sulfate overnight at 4°C and subsequently dissolved in 20 ml of 25 mM Tris–HCl pH 7.5 containing 1 mM dithiothreitol (DTT). The mixture was dialysed against the same buffer overnight at 4°C. The precipitate was pelleted and the supernatant was applied on a MonoS column (Pharmacia) equilibrated with 25 mM Tris–HCl, pH 7.5, containing 1 mM DTT. MA was separated in a linear gradient from 0 to 1 M NaCl. The fractions containing MA were separated by gel filtration on a Superose 6 column (Pharmacia) in 25 mM Tris–HCl, pH 7.5. M-PMV MA was pooled, dialysed against H_2O and 1 mM DTT at pH 6.0 and freeze-dried. For the preparation of ^{15}N-labelled material, the bacteria were grown on minimal medium with $^{15}NH_4Cl$ as the sole nitrogen source.

NMR spectroscopy

For NMR spectroscopy, ~10 mg of freeze-dried protein was dissolved in 0.6 ml of H_2O containing 10% D_2O. All NMR spectra were recorded at 500 MHz proton frequency and the temperature was maintained at 303 K throughout the experiments. The sequence-specific backbone ^1H assignments were completed using NOESY HMQC and HOHAHA HMQC spectroscopy (Marion *et al.*, 1989; Driscoll *et al.*, 1990). The secondary structure elements and ^1H side-chain assignments subsequently were determined. This procedure facilitated the interpretation of the D_2O NOESY (100 ms mixing time) spectra, which enabled the long-distance restraints to be obtained for the full structure calculation.

Structure calculation

The NOESY cross-peak intensities were measured at both 100 and 150 ms mixing times. The structures were calculated on the basis of 1092 NOE distance constraints and 45 H bond distance restraints using a dynamical simulated annealing protocol executed within the program X-PLOR (Nilges *et al.*, 1988; Brunger, 1993). The NOE restraints were composed of 248 intra-residue, 612 sequential (residue i to residues $1<i<5$) and 209 long-range (residue i to residues $i>4$) connectivities. The distance restraints were categorized into three groups on the basis of estimated NOE cross-peak intensity: strong 2.7 Å, medium 3.8 Å and weak 5.0 Å. The sum weight-averaging term was used for all distance constraints. Hydrogen-bonded NH groups were identified in helical regions by the presence of NH resonances at least 2 h after dissolving in D_2O. The helical hydrogen bonds are characterized by two restraints $r_{NH–O}$ = 2.1–2.3 Å and $r_{N–O}$ = 2.1–3.3 Å, and were introduced once they could be assigned unambiguously following initial structure calculations.

Acknowledgements

The authors acknowledge Dr Iva Pichova, Ales Zabransky and Iveta Kostalova for providing M-PMV protease. The authors would also like to thank Professor Iain Campbell and the MRC Biomedical NMR Facility at Mill Hill for access to NMR spectrometer time during the

commissioning of our own instrument. S.M. is indebted for the financial support of The Wellcome Trust and the BBSRC. T.R. and E.H. would like to also acknowledge grant VS 96074 of the Czeck Ministry of Education, grant 203/97/0416 from the Czech Grant Agency, grant CH27834 from the National Cancer Institute and grant TW00050 from the Fogarty International Centre. The co-ordinates will be deposited with the Brookhaven protein data bank and can be obtained from M.R.C. (s.conte@ic.ac.uk) and S.M. (s.j.matthews@ic.ac.uk). A comprehensive list of NMR assignments and NMR restraints is available from M.R.C. (s.conte@ic.ac.uk) and S.M. (s.j.matthews@ic.ac.uk).

References

Arnold,E. and Arnold,G.F. (1991) Human immunodeficiency virus structure—implications for antiviral design. *Adv. Virus Res.*, **39**, 1–58.

Bradac,J., Chatterjee,S. and Hunter,E. (1985) Polypeptides of Mason–Pfizer monkey virus: I. Synthesis and processing of the *gag* gene products. *Virology*, **138**, 260–275.

Brunger,A.T. (1993) *XPLOR Manual Version 3.1*. Yale University Press, New Haven, CT.

Christensen,A.M., Massiah,M.A., Turner,B., Sundquist,W.I. and Summers,M.F. (1996) Three-dimensional structure of the HTLV-II matrix protein and comparative analysis of matrix proteins from different classes of pathogenic human retroviruses. *J. Mol. Biol.*, **264**, 1117–1131.

Driscoll,P.C., Clore,G.M., Marion,D., Wingfield,P.T. and Gronenborn, A.M. (1990) Complete resonance assignment for the polypeptide backbone of interleukin-1-β using three-dimensional heteronuclear NMR spectroscopy. *Biochemistry*, **29**, 3542–3556.

Fäcke,M., Janetzko,A., Shoeman,R.L. and Kräusslich,H.-G. (1993) A large deletion in the matrix protein of the human immunodeficiency virus *gag* gene redirects virus particle assembly from the plasma membrane to endoplasmic reticulum. *J. Virol.*, **67**, 4972–4980.

Freed,E.O., Orenstein,J.M., Buckler-White,A.J. and Martin,M.A. (1994) Single amino acid changes in the human immunodeficiency virus type 1 matrix protein block virus particle production. *J. Virol.*, **68**, 5311–5320.

Gallay,P., Swingler,S., Aiken,C. and Trono,D. (1995) HIV-1 infection of non-dividing cells: C-terminal tyrosine phosphorylation of the viral matrix protein is a key regulator. *Cell*, **80**, 379–388.

Gelderblom,H.R. (1991) Assembly and morphology of HIV: potential effect of structure on viral function. *AIDS*, **5**, 617–637.

Gelderblom,H.R., Ozel,M. and Pauli,G. (1989) Morphogenesis and morphology of HIV—structure–function relations. *Arch. Virol.*, **106**, 1–13.

Hill,C.P., Worthylake,D., Bancroft,D.P., Christensen,A.M. and Sundquist,W.I. (1996) Crystal structures of the trimeric human immunodeficiency virus type 1 matrix protein: implications for membrane association and assembly. *Proc. Natl Acad. Sci. USA*, **93**, 3099–3104.

Holm,L. and Sander,C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.

Hunter,E. (1994) Macromolecular interactions in the assembly of HIV and other retroviruses. *Semin. Virol.*, **5**, 71–83.

Kliková,M., Rhee,S.S., Hunter,E. and Ruml,T. (1995) Efficient *in vivo* and *in vitro* assembly of retroviral capsids from Gag precursor proteins expressed in bacteria. *J. Virol.*, **69**, 1093–1098.

Kraulis,P.J. (1991) Molscript—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Marion,D., Driscoll,P.C., Kay,L.E., Wingfield,P.T., Bax,A. Gronenborn,A.M. and Clore,G.M. (1989) Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann–Hahn multiple quantum coherence and nuclear Overhauser–multiple quantum coherence spectroscopy: application to interleukin 1. *Biochemistry*, **28**, 6150–6156.

Marx,P.A. *et al*. (1984) Simian AIDS: isolation of a type D retrovirus and transmission of the disease. *Science*, **223**, 1083–1086.

Massiah,M.A., Staricg,M.R., Paschall,C. Summers,M.F., Christensen, A.M. and Sundquist,W.I. (1994) The three-dimensional structure of the human immunodeficiency virus type I matrix protein. *J. Mol. Biol.*, **244**, 198–223.

Matthews,S. *et al*. (1994) Structural similarity between the p17 matrix protein of HIV-1 and interferon-γ. *Nature*, **370**, 666–668.

Matthews,S., Barlow,P.N., Clark,N., Kingsman,S., Kingsman,A. and Campbell,I. (1995) The refined solution structure of the HIV-1 matrix protein, p17. *Biochem. Soc. Trans.*, **23**, 725–728.

Matthews,S., Mikhailov,M., Burny,A. and Roy,P. (1996) The solution structure of the bovine leukemia virus matrix protein and similarity with lentiviral matrix proteins. *EMBO J.*, **15**, 3267–3274.

Moscufo,N. and Chow,M. (1992) Myristate–protein interactions in poliovirus: interactions of VP4 threonine 28 contribute to structural conformation of assembly intermediates and the stability of assembled virions. *J. Virol.*, **66**, 6849–6857.

Nermut,M.V., Hockley,D.J., Jowett,J.B.M., Jones,I.A., Garreau,M. and Thomas,D. (1994) Fullerene-like organisation of HIV gag-protein shell in virus-like particles produced by recombinant baculovirus. *Virology*, **198**, 288–296.

Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association—insight from the interfacial and thermodynamic properties of hydrocarbon. *Proteins: Struct. Funct. Genet.*, **11**, 281–296.

Nilges,M. Gronenborn,A.M. and Clore,G.M. (1988) Determination of 3-dimensional structures of proteins by simulated annealing with interproton distance restraints—application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor-2. *Protein Engng*, **2**, 27–38.

Pichova,I., Zabransky,A., Kostalova,I., Hruskova,O., Andreansky,M., Hunter,E. and Ruml,T. (1997) Analysis of autoprocessing of Mason Pfizer monkey virus proteinase *in vitro*: three active forms of proteinase. In James,M.N.G. (ed.), *Structure and Function of Aspartic Proteinases: Retroviral and Cellular Enzymes*. Plenum Publishing, in press.

Rao,Z.H., Belyaev,A.S., Fry,E., Roy,P., Jones,I.M. and Stuart,D.I. (1995) Crystal structure of SIV matrix antigen and its implications for virus assembly. *Nature*, **378**, 743–747.

Rhee,S.S. and Hunter,E. (1990) A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. *Cell*, **63**, 77–86.

Rhee,S.S. and Hunter,E. (1991) Amino acid substitutions within the matrix protein of type D retroviruses affect assembly, transport and membrane association of a capsid. *EMBO J.*, **10**, 535–546.

Rhee,S.S., Hui,H. and Hunter,E. (1990) Preassembled capsids of type D retroviruses contain a signal sufficient for targeting specifically to the plasma membrane. *J. Virol.*, **64**, 3844–3852.

Smyth,M., Tate,J., Hoey,E., Lyons,C., Martin,S. and Stuart,D. (1995) Implications for viral uncoating from the structure of the bovine enterovirus. *Nature Struct. Biol.*, **2**, 22–231.

Sommerfelt,M.A., Rhee,S.S. and Hunter,E. (1992) Importance of p12 protein in Mason–Pfizer monkey virus assembly and infectivity. *J. Virol.*, **66**, 7005–7011.

Spearman,P., Wang,J., Heyden,N.V. and Ratner,L. (1994) Identification of human immunodeficiency virus type I gag protein domains essential to membrane binding and particle assembly. *J. Virol.*, **68**, 3233–3242.

Stromberg,K., Benveniste,R.E., Arthur,L.O., Rabin,H., Giddens,W.E.,Jr, Ochs,H.D., Morton,W.R. and Tsai,C.-C. (1984) Characterization of exogenous type D retrovirus from a fibroma of a macaque with simian AIDS and fibromatosis. *Science*, **224**, 289–292.

Van Der Kuyl,A., Mang,R., Dekker,J.T. and Goudsmit,J. (1997) Complete nucleotide sequence of simian endogenous type D retrovirus with intact genome organisation: evidence of ancestry to simian retrovirus and baboon endogenous virus. *J. Virol.*, **71**, 3666–3676.

Wills,J.W. and Craven,R.C. (1991) Form, function, and the use of retroviral gag proteins. *AIDS*, **5**, 639–654.

Yu,X., Yuan,X., Matzuda,Z., Lee,T.-H. and Essex,M. (1992) The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J. Virol.*, **66**, 5966–5971.

Yuan,X., Yu,X., Lee,T. and Essex,M. (1993) Mutations in the N-terminal region of human immunodeficiency virus type I matrix protein block intracellular transport of the gag precursor. *J. Virol.*, **67**, 6387–6394.

Zhou,W. and Resh,M.D. (1996) Differential membrane binding of the human immunodeficiency virus type 1 matrix protein. *J. Virol.*, **70**, 8540–8548.

Zhou,W., Parent,L.J., Wills,J.W. and Resh,M.D. (1994) Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type I gag protein which interacts with acidic phospholipids. *J. Virol.*, **68**, 2556–2569.

Received on June 4, 1997; revised on July 11, 1997