The EMBO Journal Vol.18 No.5 pp.1124–1136, 1999

Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab

Carmen Berthet-Colominas, Stéphanie Monaco, Armelle Novelli[1], Geneviève Sibaï[2], François Mallet[1] and Stephen Cusack[3]

European Molecular Laboratory Biology, Grenoble Outstation, B.P.156X, F-38042 Grenoble Cedex 9, [1]Unité Mixte 103 CNRS-bioMérieux, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cédex 07 and [2]bioMérieux, Chemin de l'Orme, 69820 Marcy l'Etoile, France

[3]Corresponding author
e-mail: cusack@embl-grenoble.fr

C.Berthet-Colominas and S.Monaco contributed equally to this work

The crystal structure of an intact molecule of HIV-1 capsid protein (p24) in complex with a monoclonal antibody fragment recognizing an epitope on the C-terminal domain has been determined at 3 Å resolution. The helical N- and C-terminal domains of p24 are linked by an extended peptide forming a flexibly linked dumb-bell-shaped molecule 75 Å in overall length. The p24 construct used is a variant with an N-terminal extension that mimics to some extent the Gag context of p24. We observed a novel head-to-tail dimer of p24 molecules which occurs through the formation of a substantial intermolecular interface between the N- and C-terminal domains. Comparison with previously observed p24 dimers shows that the same residues and secondary structural elements can partake in different interfaces revealing a remarkable stickiness and plasticity of the p24 molecule, properties which, combined with the inter-domain flexibility, are presumably important in the assembly and maturation of viral particles. Previous mutagenesis studies designed to test specific N–N and C–C homodimer interfaces do not discriminate fully against the possibility of the observed N–C interface.

Keywords: Fab/HIV-1/p24/virus assembly/X-ray crystallography

Introduction

p24 (or CA) is the major capsid protein of HIV and forms a characteristically conical shaped shell surrounding the viral RNA–nucleoprotein complex in mature virions (Figure 1; reviewed in Kräusslich, 1996; Freed, 1998). Processed HIV-1 p24 contains 231 residues but is initially expressed as part of the Gag polyprotein (p55). Gag comprises three major proteins, matrix (MA or p17), capsid (CA or p24) and nucleocapsid (NC or p7) separated by spacer peptides (p1 and p2) and terminating with the C-terminal p6 peptide (Figure 1). Virus assembly is initiated by the association of intact Gag proteins, which attach via the myristilated N-terminus of the MA to the cell inner-membrane surface. This induces budding of immature virions in which the Gag molecules are thought to extend radially inwards in paracrystalline arrays (Fuller *et al*., 1997). Only after cleavage of Gag into its distinct components by the viral protease, concomitant with budding from the host cell membrane, does p24 condense to form the mature capsid enclosing the ribonucleoprotein complex of NC with the dimeric genomic RNA. p24 is thought to play an active structural role both as part of the Gag protein and in its mature form during virus assembly, maturation and disassembly. It has also been shown that in the case of HIV-1 (but not HIV-2) a particular proline-rich loop in the N-terminal region of p24, which contains the peptide Gly89–Pro90, specifically binds cellular cyclophilin A (CypA), a peptidyl–prolyl *cis–trans* isomerase. This interaction is responsible for the sequestration of ~200 molecules of CypA into each HIV-1 particle (Franke *et al*., 1994; Thali *et al*., 1994). It seems possible that CypA is required by the virus since the presence of cyclosporine (an inhibitor of CypA) prevents the incorporation of CypA into virions, which as a result have reduced infectivity. The target of the enzyme may be CA itself, possibly to assist in the disassembly of capsids, although the evidence suggests that this is not by catalysing peptidyl–prolyl isomerization (Luban, 1996).

Although a considerable amount of literature exists on the varying effects of mutations and partial deletions of p24 on viral assembly and infectivity (e.g. Wang and Barklis, 1993; Dorfman *et al*., 1994; Reicin *et al*., 1995), it is difficult to interpret this data without detailed structural information. The aims of structural work on p24 should be to define the structure and mode of self-association (and possible association with other viral components) of p24 in both its mature and immature form. This may require the combination of high resolution X-ray and NMR structural analysis of p24 and its precursors with electron-microscopy studies of mature and immature viral particles and reconstituted capsids. This, combined with structurally based mutational analysis, with *in vitro* and *in vivo* assays of assembly and infectivity, should lead to a deeper understanding of the role of p24 and may allow definition of new targets for anti-viral agents. Considerable progress has been made in the last three years towards these goals.

The first reported structure was of the N-terminal domain of p24 (residues 1–151), obtained by NMR (Gitti *et al*., 1996). Crystals of a p24–Fab complex were reported by Prongay *et al*. (1990), and in the solved crystal structure it can be seen that the N-terminal domain (to which the Fab binds) forms dimers, but the C-terminal domain is disordered (Momany *et al*., 1996). The crystal structure of human CypA bound to the N-terminal domain of HIV-1 p24 has been solved at 2.4 Å resolution (Gamble *et al*., 1996) and reveals two distinct N-terminal domain dimer

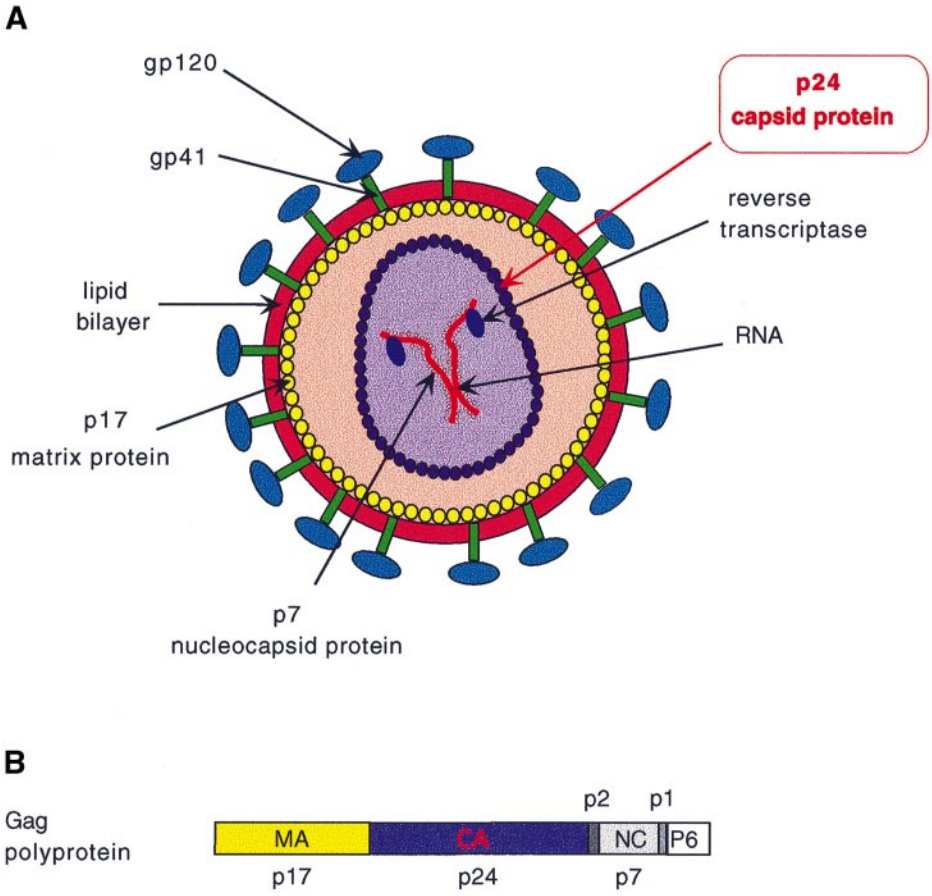


Fig. 1. Schematic diagram of the mature HIV-1 virion and the organization of the Gag protein.

interfaces. The crystal structure of the C-terminal domain of p24 has been determined more recently (Gamble *et al*., 1997). This revealed the importance of the conserved major homology region (MHR) residues in forming the core of the C-terminal domain and also showed that this domain can dimerize in two slightly different forms depending on the exact choice of the N-terminal residue (148–231 or 151–231). All the above structural studies have used the mature form of p24 starting with Pro1 or domains therefrom. The picture that emerges is that p24 comprises two flexibly linked helical domains, capable of various modes of self-association. Another important result is that creation of the N-terminus of p24 by protease cleavage, leading to the burial of the free Pro1 in the mature protein structure, is probably a key element in the maturation of the capsid. It has been hypothesized that a β-hairpin structure formed by residues 1–13 contributes to the formation of a stronger dimeric interface between the N-terminal domains in the mature protein (Gamble *et al*., 1996; von Schwedler *et al*., 1998). However, no complete structure of intact p24 has hitherto been published and the precise mode of self-association in immature and mature virions is still unclear.

In this article we present the crystal structure determined at 3.0 Å resolution of a complex of an N-terminally extended HIV-1 p24 with a monoclonal antibody Fab fragment that recognizes an epitope on the C-terminal domain. A Fab complex was chosen for crystallization since many groups have tried and failed to crystallize p24 alone, probably due to the inter-domain flexibility and the tendency towards heterogeneous oligomerization. The structure shows for the first time the complete p24 molecule and a novel mode of head-to-tail dimerization of p24 in which the N-terminal domain of one molecule interacts with the C-terminal of its 2-fold related molecule. In a second crystal form, solved at 4 Å resolution, the strict 2-fold axis relating one p24–Fab complex with another is lost and this form reveals how one N–C dimer can move relative to another via flexibility in the hinge region between the two domains.

The HIV-1 p24 used in this study is a 243-residue recombinant protein (designated RH24, derived from HIV-1 strain HXB2) with an N-terminal extension containing a His_6 tag and a slightly modified C-terminus (Cheynet *et al*., 1993). The sequence of RH24 is shown in Figure 2 compared with p24 from another HIV-1 strain (NL43) and HIV-2 (strain ROD). The N-terminal extension implies that RH24 to some extent stimulates the Gag-context of immature p24 rather than the mature form. The second component of the complex is a Fab fragment of a mouse monoclonal IgG1κ antibody (designated 13B5) which specifically recognizes an epitope on the C-terminal domain of RH24. We have also determined the structure of the Fab alone at 1.8 Å resolution and describe elsewhere the interactions of the Fab with p24 and the conformational changes upon Fab–antigen complex formation (S.Monaco, C.Berthet-Colominas, A.Novelli, N.Battaï, G.Sibaï, F.Mallet and S.Cusack, unpublished results). Antibodies against p24 are produced early after infection and their detection is widely used as a diagnostic for HIV infection

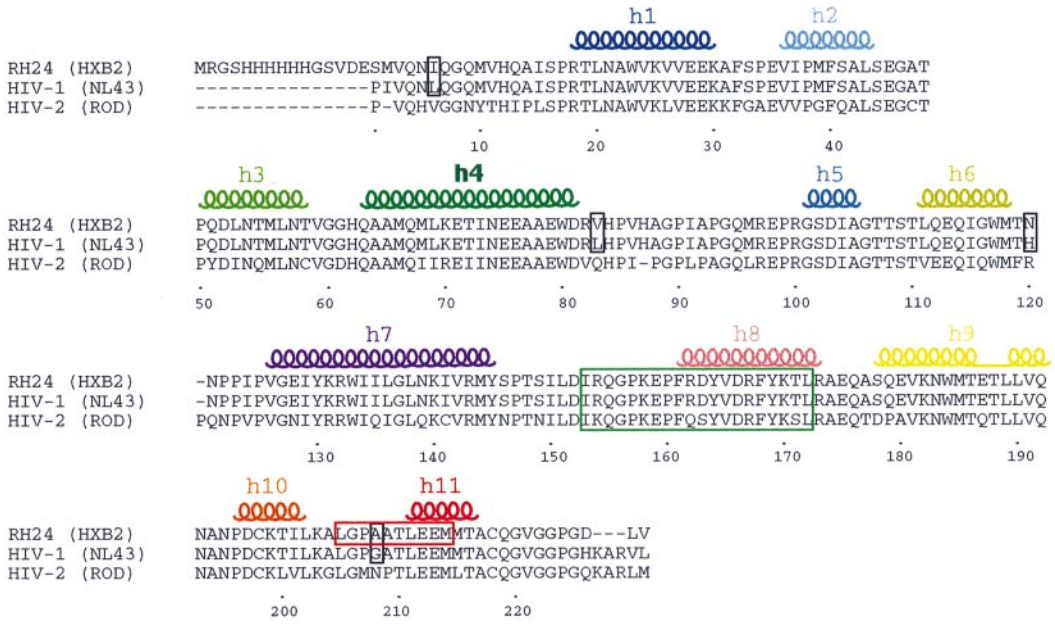


Fig. 2. Sequence alignment of RH24 (derived from HIV-1 strain HXB2) used in this study compared with mature HIV-1 (strain NL43) and HIV-2 (strain ROD). Natural point mutations of RH24 from the other HIV-1 p24 sequences are boxed in black. The major homology region (MHR) is boxed in green and the epitope of the 13B5 monoclonal antibody used in this study is boxed in red. Also shown are the positions, designations and colour code of the helices. Residue numbering is based on the RH24 sequence with residue 1 equivalent to Pro1 in the mature HIV-1 p24.

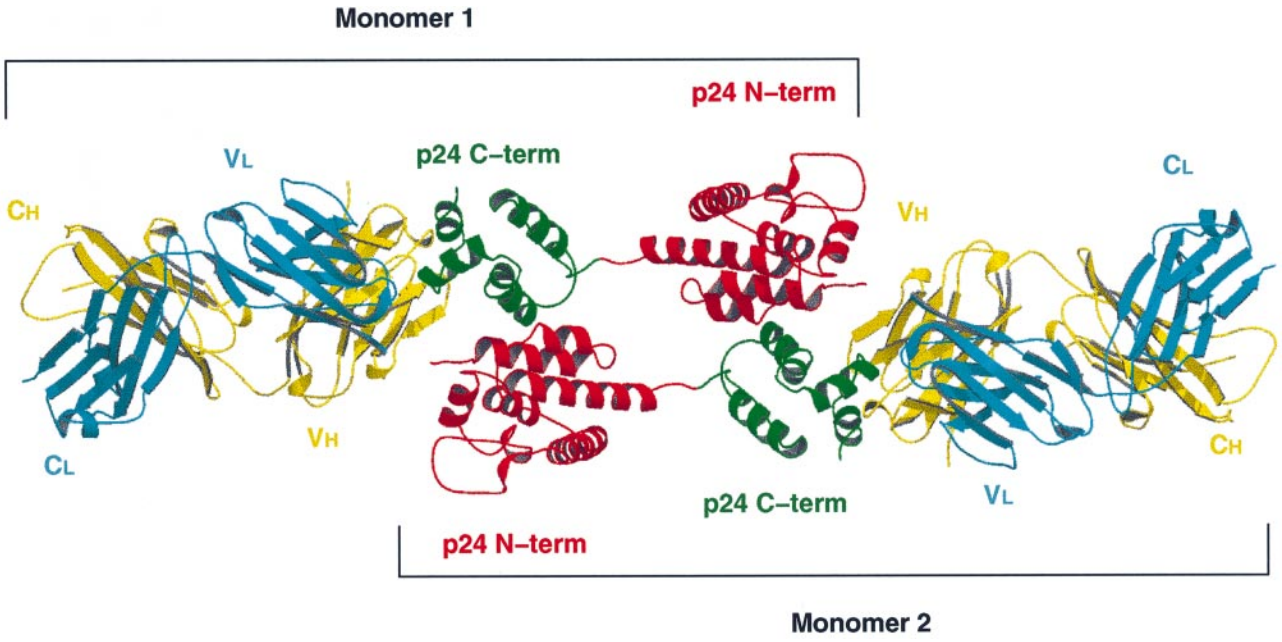


Fig. 3. Ribbon diagram of the strict 2-fold symmetrical p24–Fab dimer as observed in type II crystals with each domain labelled. The Fab light chain is blue, the heavy chain is yellow and the N- and C-terminal domains of RH24 are red and green, respectively.

(Janvier *et al.*, 1993). Although anti-p24 antibodies are not neutralizing, their continued presence in infected patients is correlated with a delayed progression to AIDS.

Results

Overall structure of p24–Fab complex

Figure 3 shows a ribbon representation of the crystallographic dimer found in the type II crystals constituted by p24–Fab complexes. In type I crystals this is a non-crystallographic dimer. Each p24–Fab complex is ~145 Å long and 35 Å wide due to the tandem longitudinal alignment of the Fab (75 Å long) and p24 (also 75 Å long), both elongated molecules. The Fab recognizes a region in the C-terminal domain of p24 in the loop between helices H10 and H11 (residues 205–213, centred around Pro207) in agreement with epitope-mapping studies using the spotscan technique (A.Novelli and L.Becquart, unpublished results). A detailed description of this interface will be given elsewhere (S.Monaco, C.Berthet-Colominas, A.Novelli, N.Battaï, G.Sibaï, F.Mallet and S.Cusack, manuscript in preparation).

The surprising result from this structure is the head-to-tail dimer of p24 resulting from an interface between the

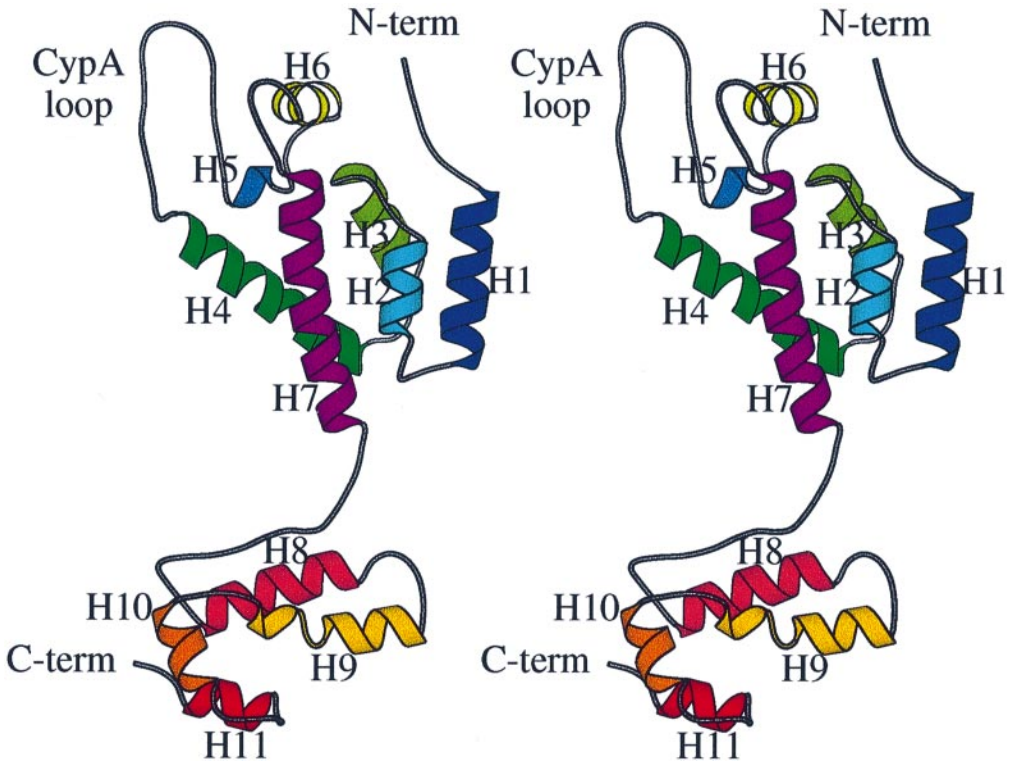


Fig. 4. Stereo diagram of the RH24 molecule showing the 11 colour-coded α-helices and with the CypA loop marked.

N-terminal domain of one molecule and the C-terminal domain of the 2-fold related molecule. This is the only p24–p24 self-contact observed in the crystal. In all previous crystallographic studies of p24, or isolated domains thereof, only homodimeric N–N or C–C interfaces have been described (Gamble *et al*., 1996, 1997; Momany *et al*., 1996).

The p24 molecule consists of two α-helical domains, the N-terminal domain (seven α-helices, H1–H7) and the C-terminal domain (four α-helices, H8–H11) interconnected by a flexible linker. A stereo view of the protein is shown in Figure 4 which introduces the colour code used in this paper for the different α-helices (the secondary structure assignment is shown in Figure 2). Both domains are similar to the structures determined previously for the individual domains (Gamble *et al*., 1996, 1997; Momany *et al*., 1996). However, there are certain significant differences discussed below with reference to Figure 5, which shows the N-terminal domain (yellow) extracted from the complex with CypA (molecule CA-1; Gamble *et al*., 1996) and the C-terminal domain (151–231, green; Gamble *et al*., 1997) superposed on the RH24 structure (red). Note that there are three natural point mutations in RH24 compared with other published p24 structures (Figure 2): Leu6→Ile (not visible in the structure), Leu83→Val, His120→Asn, Gly208→Ala, none of which has significant consequences except potentially for residue 208, which is part of the epitope.

The N-terminal domain

In the N-terminal domain of RH24 there are two disordered regions represented in white in Figure 5. The first is the cyclophilin-binding loop (residues 86–95) between helices H4 and H5 which is not unexpectedly disordered in the absence of both crystal contacts and CypA binding. The second important disordered zone is the N-terminal region preceding residue 12, where a total of 27 residues (including the His tag; Figure 2) cannot be modelled, although uninterpretable patches of electron density exist. In the mature p24 protein, the free N-terminal Pro1 produced by protease cleavage is found folded back into the protein forming a buried salt-bridge with conserved residue Asp51 (Gamble *et al*., 1996; Gitti *et al*., 1996; Momany *et al*., 1996). In some, but not all, determined structures of the N-terminal domain, residues 1–13 have been observed to form a β-hairpin which can contribute interactions to stabilize the major N-terminal dimer interface observed in the CypA complex (Gamble *et al*., 1996; von Schwedler *et al*., 1998). Disordering of the N-terminal region of RH24 is thus consistent with the idea that the buried Pro1 and β-hairpin structure can only form in the mature protein and not when p24 is in the Gag context. However, it should be remembered that both Pro1 and Ile2 are artificially substituted in the RH24 sequence (Figure 2).

Excluding the disordered regions noted above, the root mean square deviation (r.m.s.d.) for C_α positions between the N-terminal domain structure reported here and that of Momany *et al*. (1996) and Gamble *et al*. (1996; molecule CA-1) are 0.683 and 0.764 Å, respectively. The local regions where these structures differ (Figure 5) include the loop following helix H6 (residues 119–123) and that following H3 (residues 60–62). These regions are clearly flexible since they are poorly ordered or have slightly different conformations in the four distinct structures of the N-terminal domain (Momany *et al*., 1996; CA-1 and CA-2 from the CypA complex; Gamble *et al*., 1996; this work). There are also deviations at the C-terminal end of helix 7 (residues 142–145) just before the linker peptide,

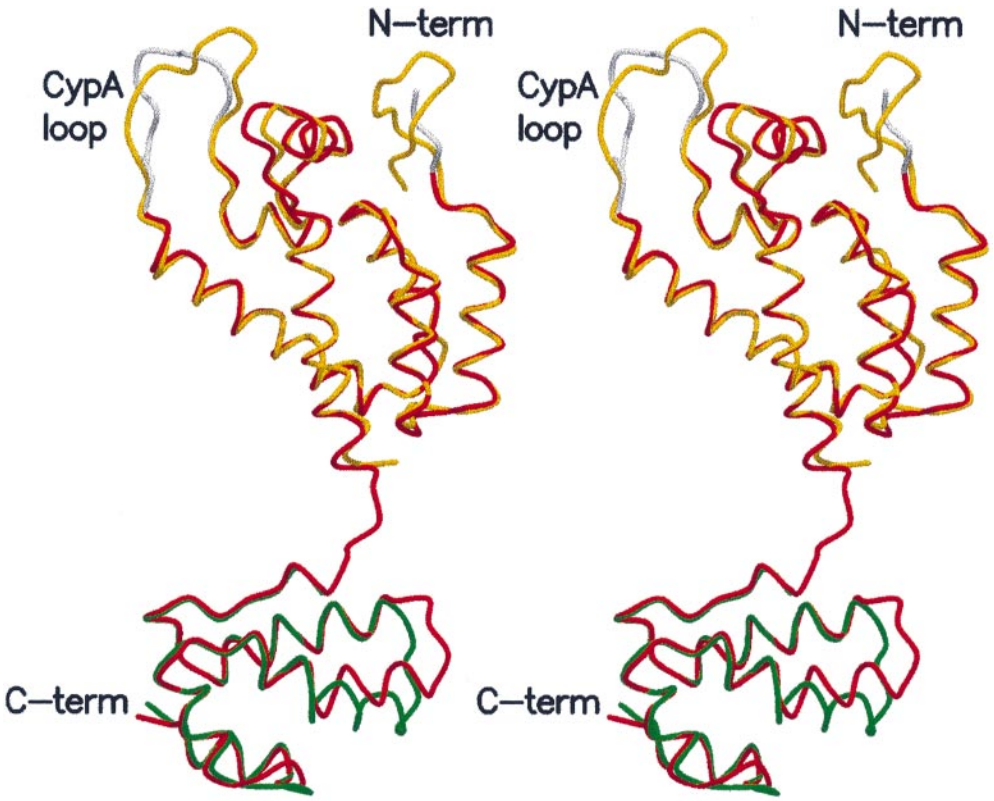


Fig. 5. Stereo diagram of the RH24 molecule (red, disordered regions white). The N-terminal domain is superimposed on the structure of CA151 extracted from the CypA complex (yellow, molecule CA-1 with the ordered N-terminal β-hairpin; Gamble *et al*., 1996). Note that there is no ordered β-hairpin structure in the second example of the CypA complex (molecule CA-2; Gamble *et al*., 1996) or in the model of Momany *et al*. (1996). The C-terminal domain is superimposed on the structure of CA146–231 (green; Gamble *et al*., 1997). The r.m.s.d. in C_α positions for superposition of the N-terminal (excluding disordered regions in white) and C-terminal domains (excluding residues 174–189) are 0.76 and 0.78 Å, respectively.

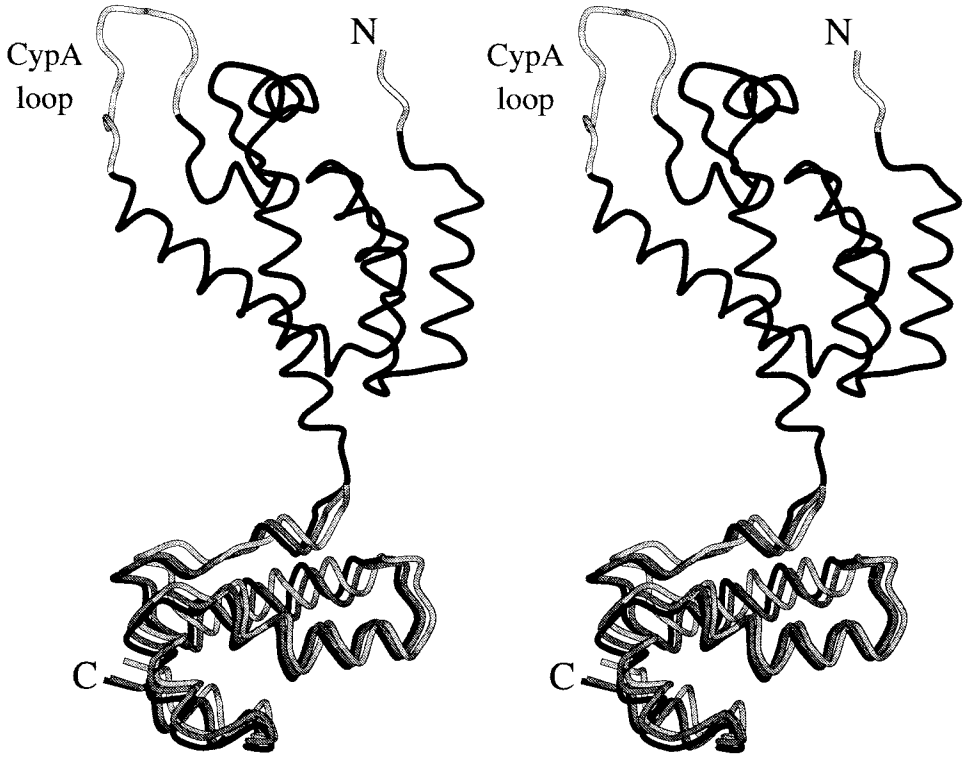


Fig. 6. Stereo diagram of the three crystallographically distinct RH24 molecules from type II crystals (black) and type I crystals (dark and light grey) after superposition of the N-terminal domains. The grey molecules were obtained from the black molecule after rigid-body refinement against the low resolution type I crystal data (see text).

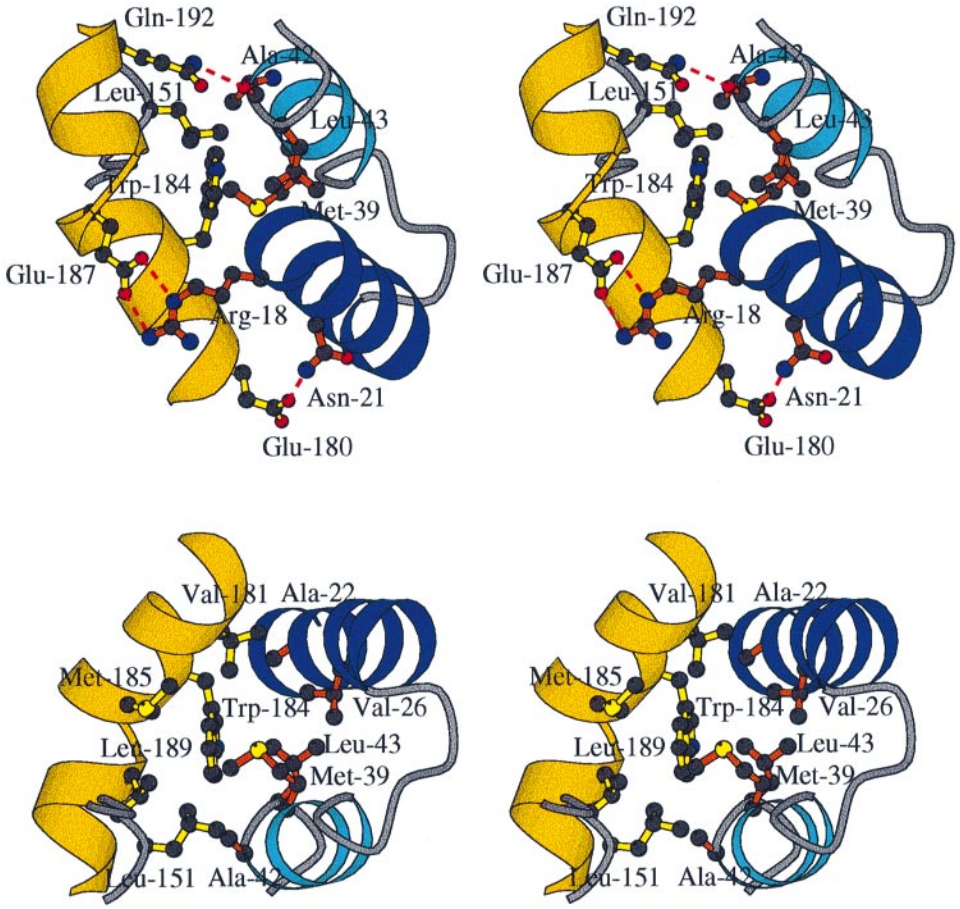


Fig. 7. Two stereo views of the N–C dimer interface. Selected residues are shown in each view for clarity. The N-terminal helices H1 and H2 are in dark blue and light blue, respectively (side-chains orange) and the distorted C-terminal helix H9 is in yellow (side-chains yellow). Putative hydrogen bonds are shown dotted in red.

Table I. Occluded accessible surfaces of each p24 monomer upon dimerization (calculated using CCP4 programme SURFACE with a 1.7 Å probe which corresponds to the van der Waals radius of water)

Protein designation or PDB code	Interface	Occluded surface (Å^2)	Reference
RH24	2× (N–C) dimer	1370	this work
1AFV	N–N dimer	1000	Momany *et al*. (1996)
1AK4	N–N dimer ('strong')	930	Gamble *et al*. (1996)
1AK4	N–N dimer ('weak')	385	Gamble *et al*. (1996)
1AM3	C–C dimer (151–231)	540	Gamble *et al*. (1997)
1AUM	C–C dimer (146–231)	585	Gamble *et al*. (1997)

which are probably related to the fact that in the CypA complex structure only the N-terminal domain up to residue 151 was present. In the RH24 complex structure, there are crystal contacts between regions 70–74 (H4) and 107–110 (between H5 and H6) of p24 and several regions of the Fab heavy chain (in the vicinity of the variable and constant domain hinge). However, these contacts do not alter the p24 conformation compared with previously published structures.

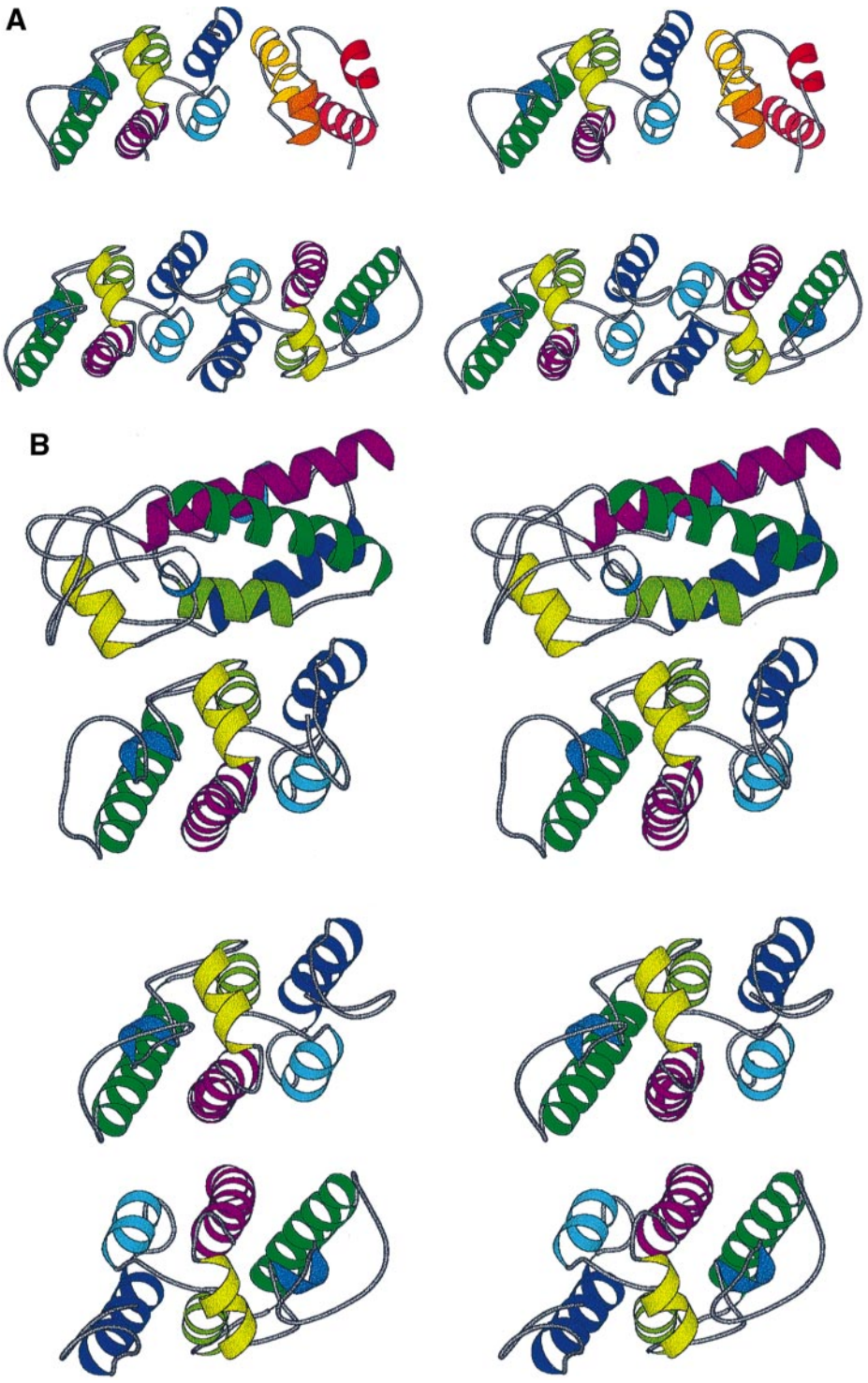
The C-terminal domain

Comparison of the C-terminal domain of RH24 and that described by Gamble *et al*. (1997) (Figure 5) shows two major differences. The most significant is in helix H9 and the preceding loop (residues 174–189). This distortion is undoubtedly due to the different dimer interface observed in the RH24 structure that involves this region of the C-terminal domain, particularly helix H9 (see below). Excluding residues 174–189 the r.m.s.d. for C_α positions between the two structures is 0.78 Å; the residues 174–189 have an r.m.s.d. of 4.4 Å. In addition, the conformation of the epitope located in the turn connecting helices H10 and H11 is also modified, probably by binding of the Fab at this position (S.Monaco, C.Berthet-Colominas, A.Novelli, N.Battaï, G.Sibaï, F.Mallet and S.Cusack, unpublished results). A minor observation is that the disulfide bond linking Cys198 and Cys218 appears to be broken in the RH24 structure (with, however, no other effect on the protein structure), which may be due to the high radiation dose given to the crystals during data collection. The MHR interactions described by Gamble *et al*. (1997), which stabilize the fold of the C-terminal domain, are unchanged.

Flexibility of the interdomain linker

Residues 146-SPTS-149 form a solvent-exposed flexible linker connecting the N- and C-terminal domains of p24, there being no interaction between the globular parts of the two domains within one molecule. Examination of omit maps in the type I crystal form suggests that the major flexibility occurs between Pro147 and Thr148. In Figure 6 the three crystallographically distinct monomers

of RH24 from the two crystal forms are superposed by means of their N-terminal domains. The shift of the relative position of the C-terminal domain demonstrates the flexibility of the interconnection region. The relative rotation is fairly small (~7.5°) but could presumably be more pronounced in the absence of crystal constraints. In terms of the head-to-tail dimer of RH24, the flexibility implies a movement of one rigid N–C unit relative to the other, without apparently altering the N–C interactions (as far as one can tell at 3–4 Å resolution). This inter-domain flexibility, combined with the intrinsic stickiness of the p24 molecule (see below) is probably responsible for the inability, so far, of anyone to crystallize intact p24. In the crystals described here, the strong interaction of the Fab with the C-terminal domain combined with that between the N- and C-terminal domains of 2-fold related p24 molecules and an additional non-specific contact between the Fab and the p24 N-terminal domain, provide the rigidity required to lock all domains into a crystal lattice. Even then, the flexibility of the linkers results in two closely related crystal forms and is probably responsible for the poor quality and twinning of the crystals. This is



a somewhat better situation than for the Fab–p24 complex crystals used by Momany *et al*. (1996), in which the Fab recognized an epitope on the N-terminal domain. In these crystals, the p24 molecules were not required for holding the crystal together and were isomorphous to crystals of the Fab alone. This led to a variable occupancy of p24, usually less than stoichiometric with the Fab, and poor crystal quality. Furthermore, there were no crystal contacts at all to the C-terminal domain which was completely disordered.

The N–C dimer interface

The unexpected result of this structure determination is the head-to-tail (N–C) nature of the p24 dimer interface, which has not been reported previously. The interaction between the two monomers is mediated predominantly by the helices H1 and H2 in the N-terminal domain and H9 in the C-terminal domain. One N–C interface has an occluded surface accessible area of 685 Å^2 giving a total of 1370 Å^2 for the dimer (calculated with a 1.7 Å probe using the programme SURFACE from the CCP4 package; Table I). Figure 7 shows two views of the interface, H1 and H2 being in dark and light blue, respectively, H9 being in yellow (same colour code as Figure 4). The interface is a typical protein–protein interface being both hydrophobic and polar in character. The hydrophobic core of the interface comprises Leu151, Val181, Trp184, Met185 and Leu189 from the C-terminal domain and Ala22, Val26 of H1 and Met39, Ala42 and Leu43 of H2 from the N-terminal domain. In addition, there are at least two hydrogen bonds, first between the side chains of Asn21 and Glu180 and secondly between the side chain of Gln192 and the main-chain carbonyl oxygen of Ala42. There is also a salt bridge between Glu187 and Arg18. Other putative interactions may exist (e.g. Lys25 interacting with Glu180) but the side-chain density is not good enough to be certain. Formation of the interface results in a severe distortion of helix H9 (Figure 5) compared with that observed by Gamble *et al*. (1997). The helix is unwound and bent in the middle due to the disruption of the normal hydrogen bonds for residues 186–187.

Comparison with other p24–p24 interfaces

Several different dimer interfaces have been found in crystals containing the capsid protein of HIV-1, which we illustrate for comparative purposes in Figure 8. The N–C interface is shown in two orientations in Figure 8A (top) and Figure 8C (top). Values for the occluded surface accessible area (calculated with a 1.7 Å probe) for each interface are given in Table I.

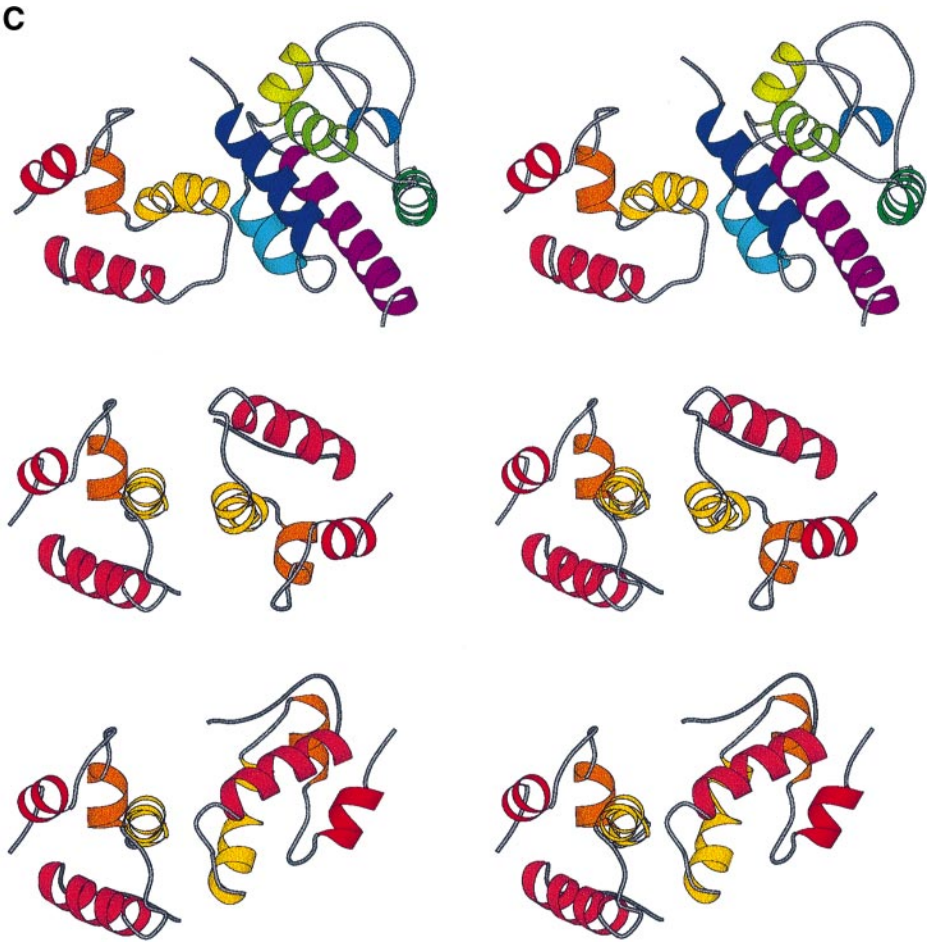


Fig. 8. Stereo views of all the reported p24 dimer interfaces using the helical colour code in Figures 1 and 3. (**A**) Top: N–C interface reported in this work. Bottom: 'strong' N–N interface from the CypA complex (interface 2; Gamble *et al*., 1996). (**B**) Top: N–N interface reported by Momany *et al*. (1996). Bottom: 'weak' N–N interface from the CypA complex (interface 1; Gamble *et al*., 1996). Note that in each diagram in (A) and (B) one of the depicted N-terminal domains is always in the same orientation to aid comparison. (**C**) Top: alternative view of N–C interface reported in this work. Middle: C–C interface from CA146–231 crystals (Gamble *et al*., 1997). Bottom: C–C interface from CA151–231 crystals (Gamble *et al*., 1997). Note that in each diagram in (C) one of the depicted C-terminal domains is always in the same orientation to aid comparison.

The N–N homodimer interface observed by Momany *et al*. (1996) is shown in Figure 8B (top). The principal contact region is a turn connecting H5 and H6 in one monomer and the corresponding turn in the 2-fold related monomer. In addition helix H1 of one monomer is in contact with helix H3 of the second monomer. The occluded surface area is 1000 Å^2 (Table I).

In the crystal containing the N-terminal domain of p24 complexed with CypA there are two different complexes in the asymmetric unit, and two different dimer interfaces have been described (Gamble *et al*., 1996). The apparently strongest (interface 2 in Gamble *et al*., 1996) is constituted by the intermolecular packing of helices H1 and H2 and the N-terminal β-hairpin (Figure 8A, bottom). The weaker interface (interface 1) is created by the helix H7 interacting as a parallel coiled-coil with its symmetry related partner (Figure 8B, bottom). The occluded surface areas are 930 and 385 Å^2 for the strong and weak interfaces, respectively (Table I).

In the crystals containing the C-terminal domain, two types of C–C homodimer interfaces are described (Gamble *et al*., 1997). In both cases the helix H9 interacts with itself through a 2-fold axis, but depending on whether the crystals contain the fragment 146–231 or 151–231 the angle between the helices differs by ~30° (Figure 8C, middle and bottom). The occluded surfaces in these interfaces are 540 and 585 Å^2 (Table I).

Consideration of all these different interfaces leads to the general conclusion that the surface of p24 is particularly 'sticky' and that this is due to surface plasticity both at the level of side-chain conformation and secondary structure (notably helix H9). Whereas several different regions of the N-terminal domain are capable of forming self-interfaces, the region of the C-terminal domain involved in interfaces is always the same (H9). It is particularly interesting to note that the regions and residues involved in N–C dimerization in this work (H1, H2 and H9) are exactly those involved in the strong interface 2 of the p24N–CypA complex (Gamble *et al*., 1996) and the C-terminal domain homodimers (Gamble *et al*., 1997). Furthermore the occluded surface areas are comparable in each case, being 1370 Å^2 for the complete head-to-tail dimer (two N–C interfaces) and 1515 Å^2 for the combined N–N (including β-hairpin) and C–C homodimers (Gamble *et al*., 1997).

Discussion

The major point of the discussion is the obvious question as to which of the many observed interfaces are biologically important during virus assembly and maturation.

The occurrence of a N–C head-to-tail dimer as observed here would appear to contradict the current view of the oligomerization of p24 both in the immature and mature virion, which is based on the assumption of only N–N and C–C interfaces (e.g. von Schwedler *et al*., 1998). Cryo-electron microscopic studies of immature HIV-1 particles and virus-like particles obtained by expression of the Gag protein in the baculovirus system have been interpreted to show that the Gag protein extends radially inwards from the inner membrane surface for ~200 Å with a paracrystalline order (Fuller *et al*., 1997). Equivalent results have been obtained with Moloney murine leukaemia virus (Yeager *et al*., 1998). Following the linear organization of the Gag protein, this would be consistent with the matrix protein being adjacent to and interacting with the membrane, followed by an unfolded segment (C-terminal of MA and N-terminal region of CA) before the two folded domains of the capsid protein, another unfolded region (p2) and finally the nucleocapsid, p1 and p6 peptides being at the lowest radius. If this radial organization is correct then one might expect any lateral

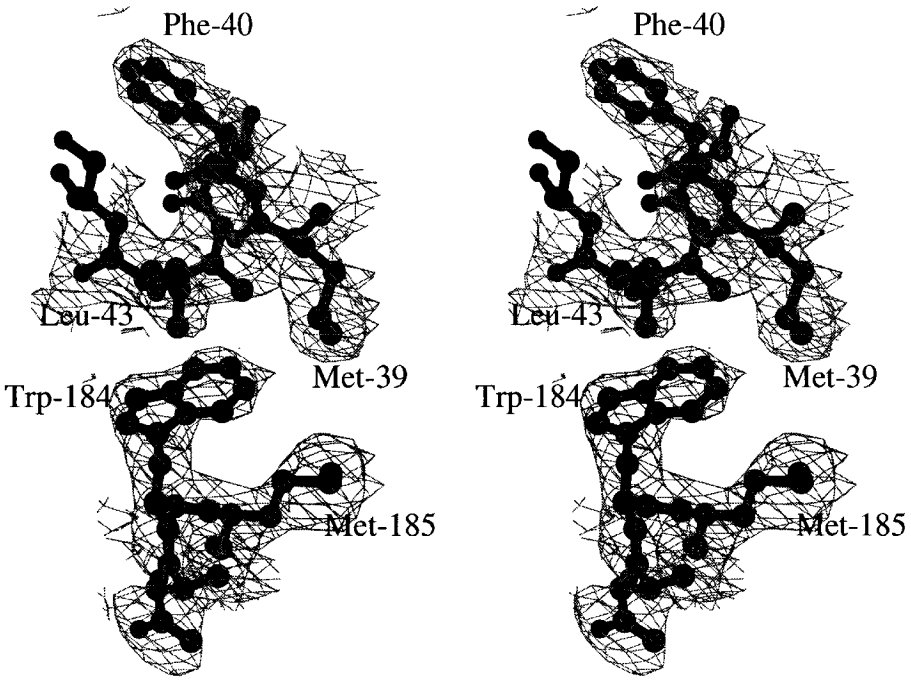


Fig. 9. Stereo diagram of $2F_o$–F_c electron density in the region of the N–C interface obtained from the final refined model and contoured at 1 σ.

contacts between the Gag proteins to again follow the linear sequence of folded domains. Oligomerization of Gag in solution has been shown to be dependent on trimerization of MA (Morikawa *et al*., 1998) and may be the initial step in the assembly of the immature virus followed by lateral interactions of CA, possibly via N–N and C–C domain interactions in a radially extended fashion. Following protease cleavage, it has been proposed that the newly released N-terminus of CA refolds to bury Pro1 at the same time creating a β-hairpin structure which reinforces and/or remodels the N-terminal domain dimerization of CA (Gamble *et al*., 1996; von Schwedler *et al*., 1998). This rearrangement would be the key step in the condensation of the mature 'conical' capsid structure. It is possible that in the mature capsid, the CA and NC proteins would also maintain their initial radial distribution. In this case, the interacting N-terminal domains of CA would form an outer shell (with the possibility of CypA fixing to its exposed binding loop on the outside), the interacting C-terminal domains would form an inner shell and filling the interior of the cone would be the NC–RNA ribonucleoprotein complex. Some supporting evidence for the double-shell model of the condensed capsid comes from recent cryo-electron microscope three-dimensional image reconstructions of *in vitro*-produced p24 cylindrical polymers, which may structurally resemble viral cones (Gross *et al*., 1997; S.Fuller and H.-G.Kräusslich, personal communication).

The preceding discussion gives a coherent and biologically reasonable picture of how p24 could interact in immature and mature virions in which there would seem to be no place for a head-to-tail dimer of the type we observe. This is largely because throughout the discussion an underlying assumption, with some supporting direct structural evidence, is that the linear sequence of protein domains in the Gag protein at all stages maps into a radial distribution in the viral particles. What then is the significance of the head-to-tail dimeric structure that we observe? First, contrary to expectation, the N–C interface may occur *in vivo* at some stage of the virus life cycle, not necessarily in stable structures. Apart from the work of Momany *et al*. (1996), this is the only high-resolution structural study where the intact p24 molecule, rather than single domains, has been used thus giving the opportunity to observe such an interface. Furthermore, this is the first study in which the p24 has had an N-terminal extension simulating the Gag-context and perhaps altering the mode of self-association. However, the saturation of two of the major interacting surfaces to form the N–C dimer does not lead to an immediate understanding of how higher order structures could be assembled. The second significance is that, even if this interface is an artefact of some kind (see below), it nevertheless reveals an unusual plasticity and stickiness of the p24 molecule, combined with inter-domain flexibility. These general properties are clearly of biological importance in permitting CA to take up different structural roles and alter its mode of self-association during viral assembly, maturation and disassembly. The third significance concerns the fact that many residues are common to both the N–C interface and the principle N–N and C–C interfaces described by Gamble *et al*. (1996, 1997). In this situation one has to be extremely careful that mutations made to test the biological importance of a particular interface really are unique to that interface. This is not the case for most of the mutations made for this purpose in the C-terminal (Gamble *et al*., 1996) and N-terminal (von Schwedler *et al*., 1998) domains (see below). We will now discuss in turn possible structural artefacts and point mutations to test interfaces.

Two particular factors (apart from the usual considerations in crystallization, such as the high protein concentration and lattice forces) could influence the type of interface observed in our crystals, first the Fab binding to the C-terminal domain and, secondly, the N-terminal extension of the RH24 molecule. We have examined whether the Fab bound to the C-terminal domain could be sterically preventing alternative dimer interfaces. If we try to create either of the two Gamble *et al*. (1997) C-terminal homodimers using our Fab–p24 complex we find that there is no steric problem in the case of the 151–231 fragment. For the 146–231 fragment an interpenetration of the Fv domain with a p24 N-terminal domain would occur which could be resolved by a significant (30°) alteration of the angle between the two p24 domains. There is no steric hinderance by the Fab to the formation of the strong CypA complex N–N interface described by Gamble *et al*. (1996) or the Momany *et al*. (1996) N–N interface. If we try to create the Gamble *et al*. (1996) weak N-terminal homodimer we find that the only steric clash is where the p24 inter-domain linkers cross each other; this could be avoided by a modest change in conformation of the linker peptide. These observations suggest that dependent on the exact range of orientational flexibility of the two p24 domains, essentially all of the previously observed interfaces could form in the presence of the Fab. We have examined this important point further by using sedimentation equilibrium experiments to measure the effect of Fab complex formation on the monomer–dimer dissociation constant of RH24 and native p24 (with Pro1). Previous measurements have reported a value of K_d = 18 μM for native p24 (Gamble *et al*., 1997; Yoo *et al*., 1997). Using a similar experimental protocol we find comparable results for both RH24 and native p24. However, under the conditions where the monomer–dimer equilibrium of RH24 (or p24) alone is readily detectable, the Fab–RH24 complex is largely monomeric (data not shown). The K_d is estimated to increase by a factor of 3–4. The Fab thus appears to weaken dimerization of RH24, perhaps due to a difficulty in making a C–C interface of the type formed by domain 146–231 as noted above. This may favour the observed N–C interface which, by inhibiting higher order oligomerization, may in turn favour crystallization.

The lack of the mature Pro1 N-terminus in RH24 prevents the formation of the β-hairpin and would weaken the N–N interface with a reduction of occluded surface area of ~30% (Gamble *et al*., 1996; von Schwedler *et al*., 1998). Under these circumstances, the observed N–C interface would probably be energetically more favourable than an N–N interface judging by occluded surface area alone. From this argument, the possible occurrence of the N–C interface in the Gag-embedded version of p24 needs to be taken seriously. The fact that an N-terminal extension radically affects the self-association properties of p24 has been shown for example by the variant MA_4–CA (four amino acids of MA fused to the N-terminus of CA, von Schwedler *et al*., 1998). *In vitro*, this protein forms

spherical particles or aggregates rather than the cylinders formed by mature CA. We have also shown that RH24 fails to form cylinders under the conditions normally used for CA polymerization (H.-G.Kräusslich, C.Berthet-Colominas, S.Monaco, A.Novelli, G.Sibaï, F.Mallet and S.Cusack, private communication). However, it should be remembered that other regions of Gag may be important in defining the self-association mode of CA in the immature virion. These include the p2 (or SP1) peptide C-terminal to CA (Accola *et al*., 1998; Wiegers *et al*., 1998), which is cleaved off in the final step of CA maturation, and possibly the p6 peptide (Garnier *et al*., 1998). To help clarify these points we are currently trying to use the same Fab to co-crystallize both mature p24 and p24 C-terminally extended with the p2 peptide.

Point mutations have been used to test the biological significance of the C–C interface (Gamble *et al*., 1997) and the principal N–N interface (von Schwedler *et al*., 1998). Assays of the effects of the mutations have been made by a variety of *in vitro* (equilibrium sedimentation, self assembly of purified proteins observed by electron microscopy) and *in vivo* (virus production, infectivity and cone formation) methods. The following mutations have been found to give rise to defective *in vitro* assembly properties, non-infectivity of virus and no mature cone formation (where tested): A22→D, E28E29→AA, M39→D, A42→D, D51→A (von Schwedler *et al*., 1998) and W184→A, M185→A (Gamble *et al*., 1997). Of these, the mutations of Ala22, Met39, Ala42, Trp184 and Met185 do not permit discrimination between the importance of N–N and C–C interactions as opposed to N–C interactions, as the residues are intimately involved in both types of interface (Figure 7). In particular, the observation (Gamble *et al*., 1997) that the W184→A and M185→A mutants are monomeric cannot be used to infer that the C–C interface is responsible for dimerization of intact p24 rather than the N–C interface. The E28E29→AA mutant is in principle somewhat more discriminatory. In the CypA complex major interface, Glu29 interacts across the N–N dimer interface with Lys30 of the other subunit, whereas Glu28 does not directly take part in the interface. In the RH24 structure, the side-chain of Glu29 is not well defined but is close to the N–C interface and again Glu28 does not directly take part in the interface. However, the double mutation E28E29→AA may considerably destabilize even the monomeric form of p24 since in the wild-type the two glutamates form part of an intramolecular charged network linking Glu28–Lys25–Glu29–Lys30–Glu35. It is in fact rather difficult to suggest mutations that would unambiguously discriminate between the two interfaces although we note that the salt bridge Arg18–Glu187 appears to be a unique feature of the N–C interface.

The crystal structure of intact p26, the capsid protein of equine infectious anaemia virus (EIAV), which is structurally homologous to HIV p24, has recently been determined (Jin *et al*., 1999). The protein used was truncated at the N-terminus and thus lacks Pro1. In two different crystal forms, a dimer is observed which is formed by yet another variant of an N–N interface. Helices 2 and 7 from each monomer form a four-helix bundle with the dimer axis perpendicular to the helices. The dimer is thus an anti-parallel arrangement of the monomers (unlike previously observed N–N dimers) with the two C-terminal domains on opposite ends. Although clearly different from the arrangement seen in the RH24 head-to-tail dimer, it is interesting to note that in both examples of structures of N-terminally modified intact capsid proteins, an anti-parallel dimer is observed.

Materials and methods

p24 and Fab purification and complex formation

HIV-1 capsid protein p24 was genetically modified by insertion of a His_6 tag at the N-terminal end as described previously (Cheynet *et al*., 1993). A consequence of this is that the proline that becomes Pro1 in the processed p24 was eliminated from the sequence together with Ile2 (Figure 2). The resulting protein, designated RH24, was over-expressed in *Escherichia coli* XLI strain, purified by metal-affinity chromatography as described (Cheynet *et al*., 1993) and stored in 50 mM sodium phosphate buffer at pH 7.4 with 5 mM dithiothreitol (DTT). Analysed by gel filtration chromatography (column Prot Pak 125 from WATERS in 0.1 M sodium phosphate buffer pH 6.8, 500 mM NaCl) in a KONTRON HPLC system, the protein was found to be >95% pure.

The monoclonal antibody 13B5 derived from immunization of mice with RH24 and directed against the C-terminal domain of p24, was provided by bioMérieux. They were papain digested and purified by protein A–Sepharose affinity chromatography and Superdex 200 gel filtration chromatography. The resulting Fab fragments were solubilized in phosphate-buffered saline (PBS) and found to be 99% pure when analysed by gel filtration chromatography.

RH24 and Fab were quantified with the Coomassie protein assay reagent (Pierce) and mixed in an equimolar ratio in the presence of 5 mM DTT. The mixture was incubated for 90 min at room temperature, dialysed in a Dialysis Cassette (Pierce) against 50 mM sodium phosphate pH 7.8 with 5 mM DTT and concentrated by centrifugation using the Centrikon system from Amicon (cut-off 30 000 *g*). The resulting complexes were analysed by non-denaturing 10% PAGE. After Coomassie Blue staining, complexes were observed to migrate as an intermediate band between RH24 and Fab without any uncomplexed RH24 or Fab. The complex was finally analysed by gel filtration and had an apparent molecular mass compatible with that of a 1:1 complex of RH24 (26.7 kDa) and Fab (45.6 kDa) and with an estimated purity of ~96%.

Crystallization

Plate-like crystals were first obtained after 3–4 days from 8% PEG 8000 at pH 8.5 using the hanging drop technique. Optimization of the crystallization conditions led to larger, but still thin, diamond shaped crystals with dimensions ~200×100×40 mm^3. They grew reproducibly at room temperature from 10% PEG 8000, 100 mM Tris–HCl pH 7.5, the protein concentration being 7 mg/ml. Almost all crystals were found to be composed of stacked platelets slightly disorientated about the plate perpendicular and thus giving severely twinned diffraction patterns. A unique untwinned crystal gave usable data to 3.8 Å resolution, the space-group being C2 with unit cell dimensions a =193 Å, b =48 Å, c = 191 Å, β =92.4° (type I crystal).

In order to improve crystal quality, different conditions were tried, which aimed to decrease the rate of growth. The crystal quality was not improved by decreasing the PEG concentration, crystallization in the presence of glycerol or growth in agarose gels. Conditions which included a high salt content were found to disassociate the complex with formation of crystals of Fab alone. Another approach was based on the analysis of the purified p24 by non-denaturing PAGE, which showed two bands that may correspond to different oxidation states of the protein. To counteract this heterogeneity, 5 mM DTT was added to the protein buffer before the complex formation as well as in the crystallization buffer and crystallization was performed at 4°C. Under these conditions crystals diffract to higher resolution but the twinning problem was not solved reproducibly. However, one good data collection was obtained on a crystal with only a small satellite, which did not deteriorate data quality. It was grown in 7% PEG 8000, 100 mM Tris–HCl pH 7.5, 5 mM DTT and 0.5 mM K_2PtCl_4 at 4°C. The platinum compound was used as a putative heavy-atom derivative but no metal was found in the resultant structure. The space group of this crystal was also C2 but with altered unit cell dimensions of a =193.1 Å, b =45.6 Å, c =132.5 Å, β=132.4° (type II crystal).

Table II. Data collection statistics for two crystal forms of the p24–Fab complex[a]

Crystal type	Type I	Type II
X-ray source	ID13 (ESRF)	ID14–EH3 (ESRF)
Detector type	Mar 300 mm Imageplate	Mar CCD
Sample temperature (K)	100	100
Wavelength (Å)	0.6889	0.945
Exposure time/image (s)	80–150	55–100
Resolution (Å)	25–4	27.7–3
Unit cell parameters (Å)	193.8, 47.9, 191.2	193.1, 45.6, 132.5
(a, b, c, β) (°)	92.4	132.4
Mosaicity (°)	0.4	0.6
Number of reflections measured	19 891	34 760
Number independent	13 023	14 742
Intensity $>3\sigma$[b] (%)	77.8 (56.8)	77.5 (26.8)
Completeness[b] (%)	72.1 (76.4)	84.7 (33.1)
Multiplicity[b]	1.9 (1.8)	2.4 (1.6)
R_{symm}[b] (%)	10.0 (17.5)	9.4 (27.6)

[a]Data were integrated with MOSFLM (Leslie, 1992).
[b]Values for the highest resolution shell in brackets.

Data collection

Data sets on two different crystal forms of the p24–Fab complex were collected (Table II). Despite screening scores of crystals, only one crystal of each type, where twinning was absent or negligible, gave usable data.

Type I. Space group C2 with unit cell dimensions, when frozen at 100 K in 30% glycerol, of a =193 Å, b =48 Å, c =191 Å, β =92.4° with two p24–Fab complexes per asymmetric unit (V_m =3.0 Å^3/Da, 60% solvent). An 80% complete data set has been measured to 4.0 Å resolution (with some data extending to 3.6 Å resolution) on a single crystal 30 μm thick on the microfocus beamline ID13 at the ESRF. The microfocus beam (30 μm diameter) may have been important in picking out less twinned regions of the crystal (Cusack *et al*., 1998).

Type II. Space group C2 with unit cell dimensions, when frozen at 100 K in 30% PEG 400, of a = 193 Å, b = 45.6 Å, c =132.5 Å, β = 132.4° with one p24–Fab complex per asymmetric unit. A 90% complete data set has been collected using a MarCCD detector on ID14-EH3 at the ESRF to 3.2 Å resolution, with some data extending to 3 Å.

The two crystal forms are very closely related. In the second form, a small rearrangement converts a non-crystallographic 2-fold axis into a crystallographic 2-fold axis with a corresponding halving of the unit cell volume.

Structure determination

Unless otherwise stated the crystallographic analysis was done with programmes from the Collaborative Computational Project 4 (CCP4) suite.

Molecular replacement. The 4 Å data obtained for the type I crystal allowed the determination of the structure by molecular replacement using the programmes AMoRe (Navaza, 1994) and XPLOR (Brunger, 1992). Three different partial models were used for the search: the Fab 13B5 structure refined to 1.8 Å resolution (S.Monaco, C.Berthet-Colominas, A.Novelli, N.Battaï, G.Sibaï, F.Mallet and S.Cusack, in preparation), the N-terminal domain of p24 corresponding to residues 1–151 (Momany *et al*., 1996) and the C-terminal domain of p24 (residues 151–220) when it became available (Gamble *et al*., 1997). Molecular replacement was tricky because two complexes were expected in the asymmetric unit, one Fab therefore being about one-third of the scattering mass with the possible complication of changes in the Fab elbow angle. In addition, the data were of mediocre quality and incomplete in the low resolution shells. AMoRe was first used to obtain rotation function solutions using the complete model of the Fab 13B5. These solutions were then refined using the Patterson Correlation (PC)-refinement algorithm of XPLOR, which permits rigid-body refinement of the four Fab domains (VH, VL, CH1, CL). It was found that the ninetieth rotation solution from the initial AmoRe run became the top solution after PC-refinement. The PC-refinement-modified Fab was subsequently used as a new search model in AMoRe to locate the first and then the second Fab molecule in the unit cell, with good packing. With these two molecules fixed, AMoRe was used to place one N-terminal domain of p24. The ninety-ninth rotation solution of the N-terminal domain gave the top translation solution and was verified by visual inspection to yield a satisfactory packing with the Fabs. The second p24 N-terminal domain was generated from the first using the non-crystallographic dimer axis defined by the two Fab solutions. After a rigid body refinement using XPLOR, a sigma A-weighted map (Read, 1986) permitted the first C-terminal domain to be orientated manually into the unbiased difference electron density. The second C-terminal domain was generated by non-crystallographic symmetry. Finally, all the different domains of the asymmetric unit (two each of VH, VL, CH1, CL, p24–N, p24–C) were refined as rigid bodies in XPLOR, the final correlation (R-factor) being 0.455 (0.458) for data between 4 and 8 Å. The type II crystal form was solved in a straightforward way by using one p24–Fab complex from the first crystal form as search model for molecular replacement with AMoRe.

Table III. Refinement statistics for Fab–p24 complex model

Resolution (Å)	15–3
Solvent content (%)	58.6
Work reflections	13 212
Test reflections	1528
R_{free}	0.284
R_{work}	0.210
Number of protein atoms (occupancy = 1)	6035
<b> Fab (Å^2)	56.0
<b> p24 (Å^2)	58.8
r.m.s. bonds (Å)	0.005
r.m.s. angles (°)	1.201
Ramachandran plot	
Favourable (%)	88.7
Additional (%)	10.0
Generous (%)	0.9 (5)
Disallowed[a] (%)	0.4 (2) Ile50L, Ala31P

Residues included: heavy chain, 1–133, 141–219 (excluding 134–140); light chain, 1–151, 156–210 (excluding 152–155); p24, 13–85, 96–145, 147–220 (excluding –12, 86–95, 146, 220+).
[a]The two residues in disallowed regions are Ile50L in a well-known classic type 1 γ-turn in the Fab light chain and Ala31P in p24.

Refinement. Refinement of the model against data between 3 and 15 Å, obtained from the type II crystal form, was performed with XPLOR v3.1 (Brunger, 1992) using a bulk solvent correction. The initial R-factors were R_{free} = 0.381, R = 0.378. In the later stages of refinement the B-factors were refined individually. Because of the relatively low resolution of the data (and corresponding high parameters/observables ratio), the free R-factor was essential to avoid over-refinement. Great care was taken not to deteriorate the geometry of the initial Fab and p24 models and for this reason only conventional positional refinement combined with manual model building was employed on the working model. However, simulated annealing runs were used to calculate omit maps and as an aid to making manual adjustments to the model. The resultant model has good R-factors (R_{free} = 0.284, R = 0.210) combined with tight restraints on geometry (r.m.s. bonds = 0.005 Å, r.m.s. angles = 1.201°). 89% of residues are in the most favourable region of the Ramachandran plot with only two residues in disallowed regions (Table III). The resultant map is as expected for this quality and resolution of data, i.e. the fold of the protein is unambiguous but side-chain electron density is often poor. Figure 9 shows a representative region of the final sigma A-weighted $2F_o$–F_c electron density map.

The refined model of the type II crystal form was used to obtain an improved model for the original type I crystal form by construction of a dimer of the p24–Fab complex and rigid-body refinement. The type I form model produced in this way has a correlation (R-factor) of 0.583 (0.399) for all measured data between 4 and 8 Å compared with corresponding values of 0.455 (0.458) for the original molecular replacement solution of the type I crystals. In the non-crystallographic dimer of the type I crystals, the pseudo 2-fold axis relating the different parts of the complex differ. The p24 N-terminal domains (p24–N) are related by an angle of 177.5°, p24–C by 170.3°, F_v (VL + VH) related by 173.5° and F_c (CH1 + CL) related by 177.7°.

Acknowledgements

We would like to thank Christine Hébrard and Philippe Pion at bioMérieux for the large-scale purification of RH24, Nadia Piga and Nicole Bataille

for kindly providing the 13B5 monoclonal antibody, and Marie-Hélène Charles and Gaspard Gervasi for helpful discussions and advice. We thank Christian Riekel, Andreas Bram and Hassan Belrhali (ESRF) for assistance with the ID13 measurements. Wim Burmeister (ESRF) for assistance on beamline ID14–EH3 and Christine Ebel (IBS) for help with the sedimentation equilibrium measurements. We are also grateful to Stephen Fuller and Rob Ruigrok for a critical reading of the manuscript and to Catherine Lawson for a preprint of the paper on the structure of EIAV p26.

References

Accola,M.A., Höglund,S. and Göttlinger,H.G. (1998) A putative α-helical structure which overlaps the capsid–p2 boundary in the HIV-1 Gag precursor is crucial for viral particle assembly. *J. Virol.*, **72**, 2072–2078.

Brünger,T.A. (1992) X-PLOR version 3.1. Yale University Press, New Haven, CT, USA.

Cheynet,V., Verrier,B. and Mallet,F. (1993) Overexpression of HIV-1 proteins in *E.coli* by a modified expression vector and their one-step purification. *Protein Exp. Purif.*, **4**, 367–372.

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Cusack,S., Belrhali,H., Bram,A., Burghammer,M., Perrakis,A. and Riekel,C. (1998) Small is beautiful: protein microcrystallography. *Nature Struct. Biol. (Synchrotron Suppl.)*, **5**, 634–637.

Dorfman,T., Bukovsky,A., Öhagen,Å., Höglund,S. and Göttlinger,H.G. (1994) Functional domains of the capsid protein of HIV-1. *J. Virol.*, **68**, 8180–8187.

Franke,E.K., Yuan,H.E.H. and Luban,J. (1994) Specific incorporation of cyclophilin A into HIV-1 virions. *Nature*, **372**, 359–362.

Freed,E.O. (1998) HIV-1 Gag proteins: diverse functions in the virus life cycle. *Virology*, **251**, 1–15.

Fuller,S.D., Wilk,T., Gowen,B.E., Kräusslich,H.-G. and Vogt,V.M. (1997) Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.*, **7**, 729–738.

Gamble,T.R., Vajdos,F.F., Yoo,S., Worthylake,D.K., Housesweart,M., Sundquist,W.I. and Hill,C.P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell*, **87**, 1285–1294.

Gamble,T.R., Yoo,S., Vajdos,F.F., Schwedler,U.K., Worthylake,D.K., Wang,H., McCutcheon,J.P., Sundquist,W.I. and Hill,C.P. (1997) Structure of the carboxyl-terminal dimerisation domain of the HIV-1 capsid protein. *Science*, **278**, 849–853.

Garnier,L., Ratner,L., Rovinski,B., Cao,S.-X. and Wills,J.W. (1998) Particle size determinants in the human immunodeficiency virus type 1 Gag protein. *J. Virol.*, **72**, 4667–4677.

Gitti,R.K., Lee,B.M., Walker,J., Summers,M.F., Yoo,S. and Sundquist, W.I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science*, **273**, 231–235.

Gross,I., Hohenberg,H. and Kräusslich,H.-G. (1997) *In vitro* assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur. J. Biochem.*, **249**, 592–600.

Janvier,B. *et al.* (1993) Prevalence and persistence of antibody titers to recombinant HIV-1 core and matrix proteins in HIV-1 infection. *J. Acquir. Immune Defic. Syndr.*, **6**, 898–903.

Jin,Z., Jin,L., Peterson,D.L. and Lawson,C.L. (1999) Model for lentivirus capsid core assembly based on crystal dimers of EIAV p26. *J. Mol. Biol.*, **286**, 83–93.

Kräusslich,H.-G. (1996) *Morphogenesis and Maturation of Retroviruses. Current Topics in Immunology.* Springer-Verlag, Berlin, Germany.

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

Leslie,A.G.W. (1992) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography, No. 26*. Daresbury Laboratory, Warrington WA4 4AD, UK.

Luban,J. (1996) Absconding with the chaperone: essential cyclophilin–Gag interaction in HIV-1 virions. *Cell*, **87**, 1157–1159.

Momany,C. *et al.* (1996) Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.*, **3**, 763–770.

Morikawa,Y., Zhang,W.-H., Hockley,D.J., Nermut,M.V. and Jones,I.M. (1998) Detection of a trimeric HIV-1 Gag intermediate is dependent on sequences in the matrix protein, p17. *J. Virol.*, **72**, 7659–7663.

Navaza,J. (1994) AmoRe: an automated package for molecular replacement. *Acta Cryst.*, **A50**, 157–163.

Prongay,A.J. *et al.* (1990) Preparation and crystallisation of a HIV p24–Fab complex. *Proc. Natl Acad. Sci. USA*, **87**, 9980–9984.

Read,R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst.*, **A42**, 140–149.

Reicin,A.S., Paik,S. and Goff,S.P. (1995) Linker insertion mutations in the HIV-1 Gag gene: effects on virion particle assembly, release and infectivity. *J. Virol.*, **69**, 642–650.

von Schwedler,U.K., Stemmler,T.L., Klishko,V.Y., Li,S., Albertine,K.H., Davis,D.R. and Sundquist,W.I. (1998) Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates vial core assembly. *EMBO J.*, **17**, 1555–1568.

Thali,M., Bukovsky,A., Komdo,E., Rosenwirth,B., Walsh,C.T., Sodroski,J. and Göttlinger,H.G. (1994) Functional association of cyclophilin A with HIV-1 virions. *Nature*, **372**, 363–365.

Wang,C.-T. and Barklis,E. (1993) Assembly, processing and infectivity of HIV-1 Gag mutants. *J. Virol.*, **67**, 4264–4273.

Wiegers,K., Rutter,G., Kottler,H., Tessmer,U., Hohenberg,H. and Kräusslich,H.-G. (1998) Sequential steps in HIV particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *J. Virol.*, **72**, 2846–2854.

Yeager,M., Wilson-Kubalek,E.M., Weiner,S.G., Brown,P.O. and Rein,A. (1998) Supramolecular organisation of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. *Proc. Natl Acad. Sci. USA*, **95**, 7299–7304.

Yoo,S., Myszka,D.G., Yeh,C., McMurray,M., Hill,C.P. and Sundquist, W.I. (1997). Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *J. Mol. Biol.*, **269**, 780–795.

Received September 8, 1998; revised December 30, 1998; accepted January 7, 1999