Expression and biochemical analysis of the entire HIV-2 gp41 ectodomain: determinants of stability map to N- and C-terminal sequences outside the 6-helix bundle core

Chan-Sien Lay[a,b], Kirilee A. Wilson[c], Bostjan Kobe[d], Bruce E. Kemp[a,e], Heidi E. Drummer[a], Pantelis Poumbourios[a,*]

[a]*St. Vincent's Institute of Medical Research, 9 Princes Street Fitzroy, Vic. 3065, Australia*
[b]*Department of Microbiology and Immunology, University of Melbourne, Vic. 3058, Australia*
[c]*Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77251-1892, USA*
[d]*Department of Biochemistry and Molecular Biology and Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia*
[e]*CSIRO Health Sciences and Nutrition, Vic., Australia*

Received 19 March 2004; revised 16 April 2004; accepted 23 April 2004

Available online 3 May 2004

Edited by Michael R. Bubb

Abstract The folding of HIV gp41 into a 6-helix bundle drives virus-cell membrane fusion. To examine the structural relationship between the 6-helix bundle core domain and other regions of gp41, we expressed in *Escherichia coli*, the entire ectodomain of HIV-2_{ST} gp41 as a soluble, trimeric maltose-binding protein (MBP)/gp41 chimera. Limiting proteolysis indicated that the Cys-591–Cys-597 disulfide-bonded region is outside a core domain comprising two peptides, Thr-529–Trp-589 and Val-604–Ser-666. A biochemical examination of MBP/gp41 chimeras encompassing these core peptides indicated that the N-terminal polar segment, 521–528, and C-terminal membrane-proximal segment, 658–666, cooperate in stabilizing the ectodomain. A functional interaction between sequences outside the gp41 core may contribute energy to membrane fusion.

Keywords: HIV; Transmembrane glycoprotein; gp41; Ectodomain; Biochemical analysis

1. Introduction

The HIV envelope glycoprotein complex, gp120–gp41, mediates virus-cell membrane fusion, which leads to viral entry. Binding by gp120 to CD4 and a chemokine receptor, CCR5 and/or CXCR4, triggers the refolding of gp41 into a trimer of hairpins, which drives membrane fusion (for review see [1–3]). gp41 is a class I fusion glycoprotein comprising an N-terminal fusion peptide that initiates membrane fusion following insertion into the target cell membrane. The fusion peptide is linked through a polar segment to an extended N-terminal α-helix that mediates glycoprotein trimerization. A disulfide-bonded region and a C-terminal α-helical segment complete an ectodomain which is anchored in the viral envelope by an ~20-residue transmembrane domain (TMD). The TMD precedes an ~150-amino acid cytoplasmic tail (for review see [3,4]).

The 6-helix bundle core domain of gp41 is composed of a central trimeric coiled coil formed by three N-terminal α-helices and three antiparallel C-terminal α-helices that pack into grooves formed on the exterior of the coiled coil [5–9]. The antiparallel packing of the C-terminal helices onto the coiled coil is considered to translocate the TMD towards the site of fusion peptide insertion, thereby juxtaposing viral and cellular membranes for merger. These interhelical interactions bury up to 3650 $Å^2$ of surface area [9], conferring stability to the 6-helix bundle and providing energy to help drive membrane fusion. The trimer-of-hairpins conformation has been observed for other retroviruses [10–12], paramyxoviruses [13,14] and a filovirus (Ebola virus) [15,16], suggesting a common fusion mechanism.

The poor solubility of gp41 when expressed in *Escherichia coli* has necessitated the use of deletion mutagenesis together with subsequent refolding and/or limited proteolysis to generate crystallizable protein [17–20]. These approaches have led to the removal of functional determinants that are located outside the 6-helix bundle core domain [5–9] such as the fusion peptide and adjacent polar segment, the conserved disulfide-bonded sequence linking the coiled coil and C-terminal helix, and the C-terminal membrane-proximal segment. More complete gp41 structures would enable a better understanding of gp41 function and would provide new targets for the development of fusion inhibitors.

We have expressed in *E. coli* the entire ectodomain of HIV-2_{ST} gp41 as a soluble, trimeric maltose-binding protein (MBP)/gp41(506–666) chimera. Limited chymotrypsin proteolysis and mass spectrometry indicated that the disulfide-bonded sequence is outside a core domain that comprises two peptides, Thr-529–Trp-589 and Val-604–Ser-666. By comparing the thermostabilities of N- and C-terminally truncated chimeras, we found that the N-terminal polar segment and C-terminal membrane-proximal segment confer stability to the gp41 ectodomain. This interaction, between sequences outside the gp41 core domain, may provide additional energy to help drive fusion.

* Corresponding author. Fax: +61-3-9416-2676.
E-mail address: apoum@ariel.its.unimelb.edu.au (P. Poumbourios).

Abbreviations: MBP, maltose-binding protein; TMD, transmembrane domain; HTLV-1, human T cell leukemia virus type 1

0014-5793/$22.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2004.04.054

2. Materials and methods

2.1. MBP/gp41 expression vectors

A modified MBP/human T cell leukemia virus type 1 (HTLV-1) gp21 vector, pMBPL-gp21(338–425) [21], was used for expression of MBP/gp41 chimeras. MBP was linked to gp41 through an Asn–Ala–Ala–Ala linker incorporating a unique *Not*I site. PCR was used to generate the HIV-2_{ST} gp41 ectodomain fragments: Gly-506–Gln-672, Gly-506–Ser-666, Ala-521–Ser-666, Ala-522–Ser-654, and Thr-529–Val-657, using pJSP4/27-H6 [22] as the template. The forward primers, 5′-CCCGCGAGGGCGGCCGCG**GGT**GTATTCGTGCTA, 5′-ATAAGAATGCG**GCC**GCGATGGGCGCGGCGTCCTTGACG, 5′-C ACGACAGCGGCCGCT**GCA**ATGGGCGCGGCG and 5′-CTGGGCGACGCGGCCGCG**ACG**CTGTCGGCTCAG, incorporate the *Not*I site (underlined) and the N-terminal gp41 residues, Gly-506, Ala-521, Ala-522, and Thr-529, respectively (bold). The reverse primers, 5′-GCCAAAGTCGAC*TTA***GCT**ATTTAATTTTTGTAG, 5′-ATGGCCGACGTCGAC*TTA***AAC**ATCCCAGCTATTTAATTTTTG, 5′-CTGAATATAGTCGAC*TTA***GGA**GGTTAAATCAAACCA and 5′-CCGGCCCTAGTCGAC*TTA***CTG**AATATATTTGATCCAG, incorporate the C-terminal gp41 amino acids Ser-654, Val-657, Ser-666 and Gln-672, respectively (bold), a stop codon (italics) and a *Sal*I site (underlined). The PCR products were ligated into the MBP expression vector through *Not*I–*Sal*I. DNA sequences were confirmed using ABI prism BigDye terminator (Applied Biosystems, Foster City, CA).

2.2. Expression and purification of MBP/gp41 chimeras

MBP/gp41 chimeras were induced at room temperature using 0.2 mM isopropyl-β-D-thiogalactopyranoside in *E. coli* strain BL21(DE3). After 18 h, the pelleted cells were lysed by sonication in ice-cold MES buffer (100 mM 2-[*N*-Morpholino]ethanesulfonic acid, pH 5, and 300 mM NaCl, and 1 mM EDTA) and clarified by centrifugation (1 h at $20\,000 \times g$). MBP/gp41 was affinity-purified using amylose–agarose as recommended (New England Biolabs). MBP/gp41 oligomers were further purified by gel-filtration chromatography using Superdex 200 (HiLoad 26/60) equilibrated in MES buffer. Gel-filtration experiments were calibrated with blue dextran, ferritin (440 kDa), aldolase (158 kDa) and ovalbumin (43 kDa) (Amersham Biosciences).

2.3. Analytical ultracentrifugation

MBP/gp41(522–654) oligomers were subjected to sedimentation equilibrium on a Beckman XL-A analytical ultracentrifuge equipped with Absorbance Optics, an An-60Ti rotor and filled-epon centerpieces (12 mm path length). MBP/gp41(522–654) oligomers at 0.8, 0.4 and 0.08 mg/ml in MES buffer, pH 6.5, were centrifuged at 20 °C for 16 h, at 12 000 rpm. The final equilibrium distributions were determined from absorption measurements at 205, 280 and 295 nm. Theoretical partial specific volume calculated from the amino acid composition of HIV-2_{ST} MBP/gp41(522–654) was 0.743 and the solvent density of the MES buffer was calculated to be 1.013 g ml^{-1} at 20 °C. To calculate molecular weight, data were fitted using the SEDEQ1B program (provided by Allen Minton, NIH, Bethesda, MD) to a single species model. No systematic deviation of the residuals was evident in any concentration set. The optical density at 280 or 295 nm as a function of radial distance was fitted by global analysis assuming a single solute.

2.4. In vitro oxidation

MBP/gp41 trimers (0.09 mg/ml) were oxidized using 1.5 mM copper II, 4.5 mM phenanthroline in 100 mM Tris, pH 8.5, for 18 h at room temperature. The oxidized protein was repurified using a Superdex 200 (HiLoad 26/60) gel-filtration column equilibrated in MES buffer or 50 mM glycine, pH 6.5, and 300 mM NaCl. The presence or absence of free sulfhydryls in MBP/gp41 was determined by treating with the alkylating agent, 1,4-vinylpyridine (Sigma) followed by electrospray mass spectrometry as described [21,23].

2.5. Limited chymotrypsin proteolysis of MBP/gp41 chimeras

MBP/gp41(506–666) trimer (0.5 mg/ml) was proteolyzed with sequencing-grade chymotrypsin (Roche) at a 1:50 ratio of protease to protein (w/w) in 100 mM Tris, pH 8.5. Proteolysis was quenched by the addition of 0.1% trifluoroacetic acid prior to mass spectrometry.

2.6. Electrospray mass spectrometry

Methanol/chloroform-precipitated proteins were resuspended in 15% (v/v) acetic acid, 30% (v/v) acetonitrile and directly infused into the mass spectrometer using a nanoelectrospray ion source (PE Sciex API III instrument) and NanoES spray capillaries (Protana Engineering A/S). Ion spectra were visualized with Tune 2.5-FPU software and deconvoluted by hypermass reconstruction using MacSpec 3.3 (PE Sciex).

2.7. Thermal aggregation assays

The thermostability of MBP/gp41 trimers was assessed using a thermal aggregation assay [23]. Briefly, purified MBP/gp41 trimers were exchanged into 50 mM glycine, pH 6.5, 300 mM NaCl, using a Superdex 200 HR (10/30) column. The trimers were heat-treated at 48, 49 or 50 °C for 5 min, cooled on ice, and then centrifuged at $16\,000 \times g$ for 4 min prior to gel filtration on a Superdex 200 HR (10/30) column at a flow rate of 0.5 ml min^{-1}.

3. Results

3.1. Use of MBP as a solubilization partner for the entire ectodomain of HIV-2_{ST} gp41

We investigated whether the MBP expression system could be used for the bacterial expression of the HIV gp41 ectodomain in soluble form. We prepared a chimera composed of MBP linked through a trialanine linker to an HIV-2_{ST} gp41

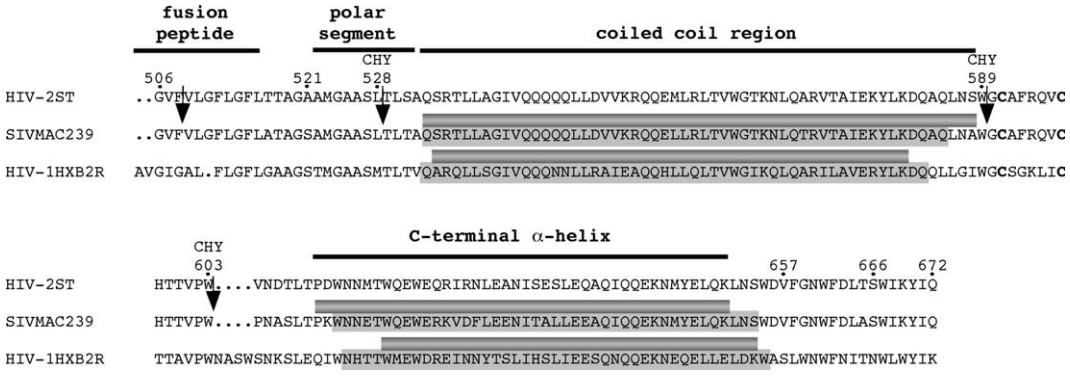

Fig. 1. Alignment of gp41 ectodomain sequences from HIV-2_{ST}, SIV_{mac239} and HIV-1_{HXB2R} and the organization of functional elements (HIV-2_{ST} numbering is used). The N- and C-terminal α-helices of the SIV_{mac239} [9] and HIV-1_{HXB2R} [8] 6-helix bundles are indicated as dark gray cylinders. The residues encompassed by the protease-resistant core of SIV_{mac239} [17] and HIV-1_{HXB2R} [18] gp41 are highlighted in light gray. The conserved Cys residues forming the intramolecular disulfide are shown in bold type. The chymotrypsin cleavage sites identified in this study by mass spectrometry for MBP/gp41(506–666) are indicated by arrows and labeled CHY.

ectodomain fragment, Ala-522–Ser-654, MBP/gp41(522–654) (Fig. 1). Purified MBP/gp41(522–654) eluted as a symmetrical peak with Superdex 200 gel filtration (Fig. 2A). The trimeric state of this protein was confirmed by sedimentation equilibrium ultracentrifugation, the experimental data closely fitting the slope of a theoretical trimer (168 kDa) with an observed molecular mass of 174 ± 5.2 kDa (Fig. 2B). MBP/gp41 chimeras incorporating homologous HIV-1 (YU2 and BH8 strains), HIV-2 (ROD) and SIV (MAC239) sequences gave lower yields of protein of heterogeneous oligomeric structure (data not shown). We, therefore, used the HIV-2_{ST} sequence to model the gp41 ectodomain.

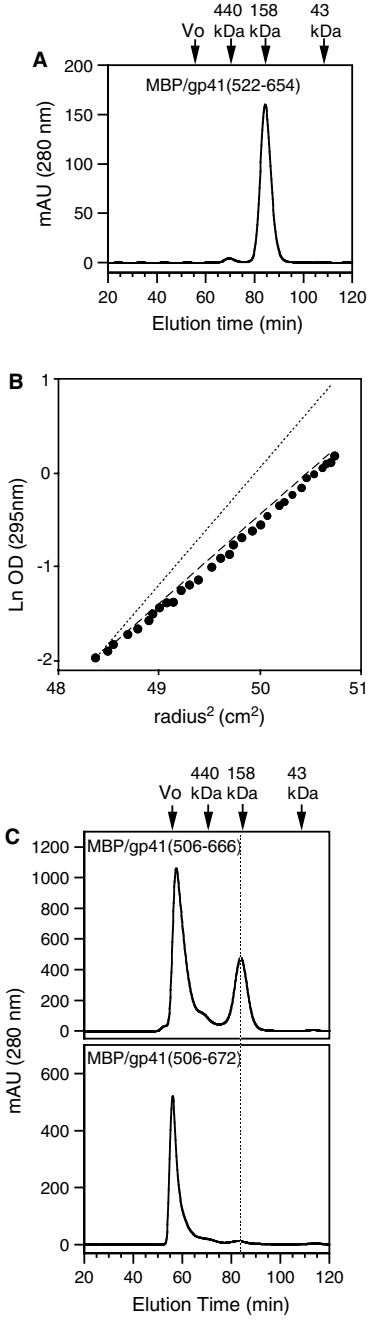

Fig. 2. (A) Superdex 200 (26/60) gel-filtration profile of MBP/gp41(522–654) oligomer. The column was calibrated using ferritin (440 kDa), aldolase (158 kDa) and ovalbumin (43 kDa) (arrows). Vo, void volume. Flow rate: 2 ml min^{-1}. (B) Sedimentation equilibrium profile of MBP/gp41(522–654) oligomer. The log of the optical density at 295 nm as a function of the square of the radial distance is shown (closed circles). The lines of best fit for theoretical tetramer (dotted line) and trimer (dashed line). No systematic deviation from the residuals was evident. (C) Superdex 200 (26/60) gel-filtration profiles of MBP/gp41(506–666) and MBP/(506–672) chimeras following elution from amylose agarose. Flow rate: 2 ml min^{-1}. Dashed line indicates the elution position of trimeric MBP/gp41(522–654).

We next determined if MBP could be used as a solubilization partner for the entire HIV-2_{ST} gp41 ectodomain, including the fusion peptide. MBP/gp41(506–666) eluted as two peaks with Superdex 200 gel filtration: a high-molecular weight aggregate and a minor species that co-eluted with MBP/gp41(522–654) trimer (Fig. 2C). Extension of the C terminus by six residues in MBP/gp41(506–672) resulted in aggregated protein. These data are consistent with the HIV-2_{ST} gp41 ectodomain terminating at Ser-666 [24,25], with residues 667–672 likely to represent a portion of the TMD.

3.2. Oxidation of the gp41 disulfide bond in vitro

The gp41 ectodomain contains two conserved cysteines, Cys-591 and Cys-597, forming an intramolecular disulfide bond that is required for correct Env glycoprotein folding and function [26]. We assessed the redox state of MBP/gp41(506–666) by using mass spectrometry to monitor the alkylation of free sulfhydryls by 1,4-vinylpyridine. The treatment of MBP/gp41(506–666) with 1,4-vinylpyridine led to the addition of 214 Da, consistent with both cysteines being reduced (Figs. 3A and B). Disulfide bond formation was induced using copper II[1,10 phenanthroline]$_3$-catalyzed oxidation in vitro. The mass of oxidized MBP/gp41(506–666) trimer was not altered by 1,4-vinylpyridine treatment, consistent with quantitative disulfide bond formation (Fig. 3C). The trimeric structure of MBP/gp41(506–666) was retained following the oxidation procedure as indicated by gel-filtration chromatography (data not shown). Approximately 5% of the purified oxidized trimer migrated as high-molecular weight disulfide-linked oligomers in SDS–PAGE under non-reducing conditions (Fig. 3D, asterisk) indicating that the yield of intrachain disulfide formation was approximately 95%. This procedure yielded approximately 0.3 mg oxidized MBP/gp41(506–666) trimer per litre of bacterial culture.

3.3. Limited chymotrypsin proteolysis of MBP/gp41(506–666)

The availability of MBP/gp41(506–666) now allows an assessment of the gp41 core domain by limited proteolysis in the context of an intact, trimeric ectodomain. Oxidized MBP/gp41(506–666) trimers were treated with chymotrypsin (protease:protein ratio = 1:50 [w/w]) at 37 °C for various times. Mass spectrometry of a 30-s incubation revealed two major proteolytic products, MBP/gp41(G506-W589) (49 685 Da) and MBP/gp41(G506-W603) (51 276 Da) (Fig 4A); a peak corresponding to Gly-590–Trp-603 was not observed. Thus, chymotrypsin cleaves rapidly on either side of the Cys-591–Cys-597 disulfide indicating that this sequence is exposed and likely forms a poorly ordered structure outside the core domain. After 60 min of incubation, the largest gp41 peptides detected were fragments of 7039 Da, corresponding to gp41(T529-W589) (Fig 4B), and 7679 Da, corresponding to gp41(V604-S666) (Fig 4C). Although we observed MBP/gp41(506–508) (40 814 Da) at the 60 min time point, a species corresponding to gp41(V509-L528) was not detected, perhaps due to poor ionization of this peptide (data not shown). These data

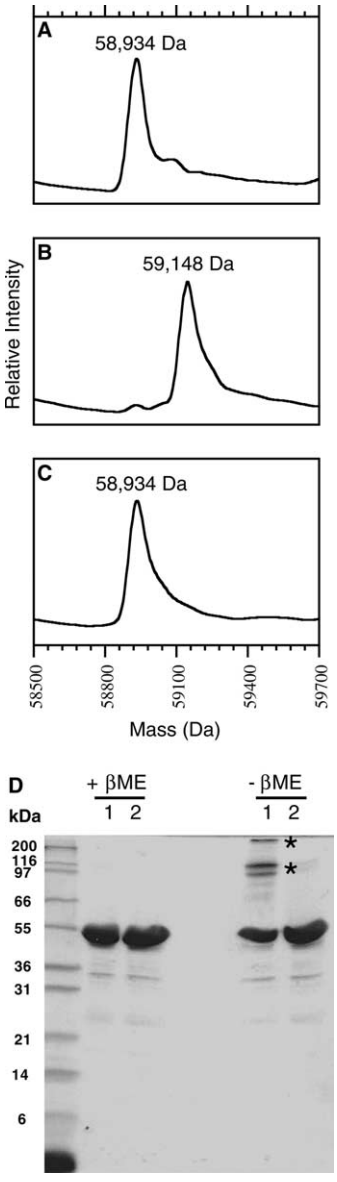

Fig. 3. (A) Electrospray mass spectrometry of MBP/gp41(506–666) trimer; (B) alkylation of MBP/gp41(506–666) trimer with 1,4-vinylpyridine adds 214 Da, consistent with the presence of two free sulfhydryls; (C) in vitro-oxidized MBP/gp41(506–666) is not modified by treatment with 1,4-vinylpyridine, consistent with disulfide formation; (D) SDS–PAGE under reducing (+ βME) and non-reducing (− βME) conditions of MBP/gp41(506–666) after (lanes 1) and before (lanes 2) oxidation in vitro.

indicate that the core domain of gp41 comprises the N- and C-terminal α-helical segments gp41(T529-W589) and gp41(V604-S666), respectively. These peptide components of the putative core domain obtained from soluble HIV-2_{ST} gp41 ectodomain are larger than previously determined by limited proteolysis of insoluble HIV-1 and SIV gp41 constructs (Fig. 1) [17–19].

3.4. *N- and C-terminal sequences external to the 6-helix bundle domain confer stability to the gp41 ectodomain*

We next constructed the MBP/gp41(529–666) chimera based on the assumption that Thr-529–Ser-666 encompasses the gp41 core. Unexpectedly, this protein eluted as a high-molecular weight aggregate with little evidence of trimer in gel-filtration

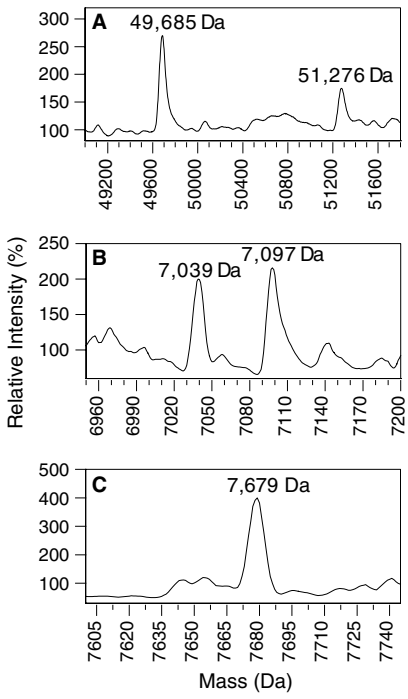

Fig. 4. Electrospray mass spectrometry of MBP/gp41(506–666) digested with chymotrypsin at 37 °C for 30 s, Panel A and 60 min, Panels B and C. Panel A, 49 685 and 51 276 Da species correspond to MBP/gp41(506–589) and MBP/gp41(506–603), respectively; Panel B, 7039 Da species corresponds to gp41(529–589); Panel C, 7679 Da species corresponds to gp41(604–666). The 7097 Da species observed in Panel B does not correspond to a theoretical chymotryptic peptide of gp41 nor MBP. This species was also observed at the 30 s time point and may be derived from a contaminant.

experiments (Fig. 5). Trimerization ability was restored by a 9-residue C-terminal truncation [chimera MBP/gp41(529–657)] or by extending the N-terminal residues from Thr-528 to Ala-521 [chimera MBP/gp41(521–666)]. These data indicate that residues 521–528, comprising the polar linker between the fusion peptide and coiled coil [27], are required for the stable accommodation of the membrane-proximal residues, 658–666, within a trimeric chimera.

We next compared the stability of in-vitro-oxidized MBP/gp41(529–657), MBP/gp41(521–666) and MBP/gp41(506–666) trimers by using a thermal aggregation assay where the temperature-dependent conversion of trimer to high-molecular weight aggregate is monitored by gel-filtration chromatography [23]. The trimeric structure of MBP/gp41(529–657) was retained after a 5-min heat treatment at 48 °C, whereas ~80% of this protein formed aggregates after treatment at 50 °C (Fig. 6). The heat treatment may cause the aggregation of gp41 trimers through the exposure of hydrophobic regions. Alternatively, thermal aggregation may occur through association of denatured gp41 domains (MBP retains its monomeric structure even after treatment at 55 °C for 5 min, data not shown). By contrast, MBP/gp41(521–666) resisted aggregation with >95% of this protein retaining its trimeric structure following the 50 °C treatment. Inclusion of the fusion peptide in MBP/gp41(506–666) led to increasing amounts of aggregate forming after 48 and 49 °C treatments, respectively (Fig. 6), and the 50 °C treatment resulting in total protein precipitation (data not shown). Thus, the N-terminal polar linker 521–528

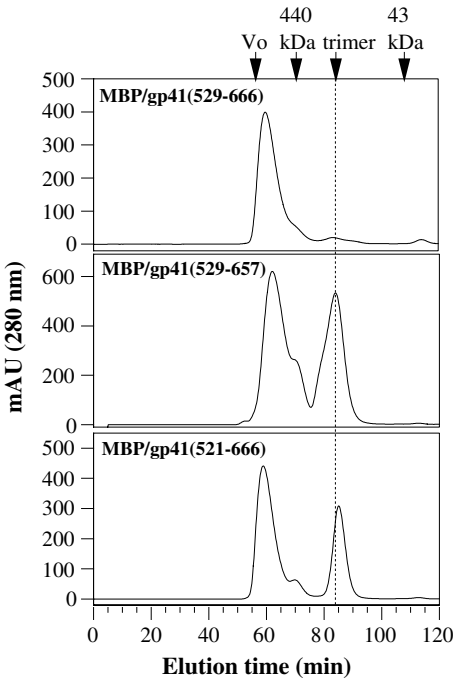

Fig. 5. Superdex 200 (26/60) gel-filtration profiles of MBP/gp41(529–666), MBP/gp41(529–657), and MBP/gp41(521–666) after elution from amylose–agarose. Dashed line indicates the elution position of trimeric MBP/gp41(522–654). Flow rate: 2 ml min^{-1}.

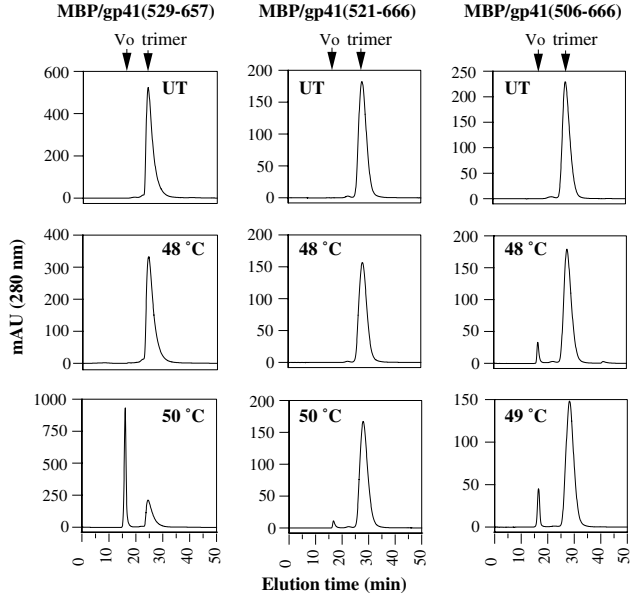

Fig. 6. Thermal aggregation-gel-filtration analysis of MBP/gp41 chimeras. The effects of a 5 min treatment at the indicated temperatures on the trimeric structures of in vitro oxidized MBP/gp41(529–657), MBP/gp41(521–666) and MBP/gp41(506–666) were monitored by Superdex 200 HR (10/30) gel filtration at a flow rate of 0.5 ml min^{-1}. The treatment of MBP/gp41(506–666) for 5 min at 50 °C resulted in precipitation of the total protein.

and C-terminal membrane-proximal sequence 658–666, which are located outside the 6-helix bundle core domain, confer stability to the gp41 ectodomain against exposure of hydrophobic sequences. However, the fusion peptide has a destabilizing influence on the ectodomain, consistent with the observations that hydrated class I fusion peptides are unstable [28].

4. Discussion

We have shown that the entire ectodomain of HIV-2 gp41 (Gly-506–Ser-666) can be obtained as a soluble trimer when MBP is used as an N-terminal fusion partner. MBP may have general utility for the soluble expression of retroviral fusion glycoproteins as we have also used this approach to express the entire HTLV-1 gp21 ectodomain in soluble form [23]. This capacity to obtain the entire gp41 ectodomain enabled for the first time, a biochemical assessment of the relationship between the gp41 core domain and structural elements located outside the core.

The antiparallel N- and C-terminal α-helical core segments of retroviral TMs are linked through the disulfide-bonded sequence. We found that the gp41 disulfide does not form spontaneously in *E. coli*, but its formation can be induced by oxidation in vitro with copper II [1,10 phenanthroline]$_3$. Despite the presence of the disulfide in MBP/gp41(506–666), limited proteolysis indicated that this region is exposed and possibly poorly ordered compared to the core domain. In contrast, the corresponding region in other retroviruses, including HTLV-1 and murine leukemia virus, forms a protease-resistant structure mediating a chain reversal that is stabilized by a polar interaction between the disulfide-bonded loop and the base of the coiled coil [10,11,23,29]. Furthermore, the disulfide-bonded region of gp41 is dispensable for 6-helix bundle formation [6,8], whereas the polar interaction between the disulfide-bonded region and the coiled coil is an essential component of the HTLV-1 trimer of hairpins core domain [30]. These structural differences may be reflected in functional differences observed for this region in HIV-1 and HTLV-1 TMs, the gp41 disulfide-bonded region appearing to form an adaptable structure important for gp120 association whereas its gp21 counterpart mediates fusion activation [30–32].

The protease-resistant core domain identified in this study, Thr-529–Trp-589 and Val-604–Ser-666, includes an additional 4 N-terminal and 11 C-terminal residues, respectively, when compared to previous investigations employing proteolysis of insoluble protein lacking the fusion peptide, polar linker and membrane-proximal regions (see Fig. 1) [17,18]. These additional N- and C-terminal sequences may be structured as they are resistant to chymotrypsin despite the presence of several protease targets (Leu-530, Trp-655, Phe-658, Trp-661 and Leu-664). The additional sequences encompass structural and functional elements that have been observed in other retroviruses. For example, the analogous N-terminal residues of HTLV-1 gp21 form a functionally important N cap that terminates the coiled-coil forming helices [11,30], while the C-terminal residues encompass a tryptophan-rich motif shown to be important for HIV-1 glycoprotein fusogenicity [33,34].

The MBP/gp41(529–666) chimera, based on the protease-resistant core, was obtained as a high-molecular weight aggregate from *E. coli*. However, an 8-residue N-terminal extension of the gp41 domain, or a 9-residue C-terminal truncation restored trimerization. The N-terminal polar linking segment, 521–528, is thus required for the accommodation

of the membrane-proximal segment, 658–666. The C-terminal segment, 658–666, may promote aggregation of trimers, unless shielded by the polar segment, 521–528. MBP/gp41(521–666) containing these terminal sequences exhibited greater thermostability than MBP/gp21(529–657), lacking these elements. These terminal regions may interact functionally or directly to stabilize the ectodomain. Munoz-Barroso et al. [34] showed that a cluster of aromatic residues within the membrane-proximal segment of HIV-1 gp41 is important to drive fusion beyond the pore formation phase. An interaction between hydrophobic residues within the N-terminal polar linking segment and the membrane-proximal aromatic cluster may provide energy for pore expansion and the completion of fusion. We have observed conserved functional roles for these regions in the HTLV-1 retrovirus. Sequential C-terminal truncation of full-length MBP/gp21 ectodomain is associated with incremental decreases in trimer of hairpins stability [23], while an I334A substitution within the polar linking region both decreases MBP/gp21 stability and blocks the pore expansion phase of fusion (Wilson et al., 2004; submitted). Functionally important interactions between membrane-proximal sequences extend to the *Orthomyxoviridae* family of viruses, which also utilize a class I fusion mechanism. The tight binding between a C-terminal extended sequence and the N-terminal end of the coiled coil, including the N cap, of influenza virus HA2 promotes the initial lipid-mixing phase of fusion [35].

The structures of the HIV-1_{HXB2R} and SIV_{MAC239} 6-helix bundles superimpose with a root mean square deviation of less than 1 Å [9], while sharing 53.5% amino acid sequence identity. The HIV-2_{ST} gp41 core shares 51.3% and 86.8% amino acid sequence identity, respectively, with HIV-1_{HXB2R} and SIV_{MAC239}, indicating that their structures are likely to be very similar. The N- and C-terminal sequences of HIV-2_{ST} gp41, found in this study to be important for ectodomain stability, are even more closely related (57.6% identity) to their HIV-1 counterparts than are the 6-helix bundle cores. Our findings for HIV-2 gp41 may, therefore, also apply to HIV-1 gp41. The novel recombinant forms of HIV-2 gp41 identified in this study may facilitate future structural investigations of these N- and C-terminal functional regions to provide a more complete understanding of viral fusion.

Acknowledgements: We thank G.J. Howlett for help with sedimentation equilibrium ultracentrifugation. pJSP4/27-H6 was obtained from B. Hahn and G. Shaw through the NIH AIDS Research and Reference Reagent Program. This work was supported by NHMRC Grants 156714, 205306 and 9937804. B.K. is an NHMRC Senior Research Fellow. B.E.K. is an Honorary NHMRC Fellow and A.R.C. Federation Fellow at CSIRO Health Sciences and Nutrition.

References

[1] Reeves, J.D. and Doms, R.W. (2002) J. Gen. Virol. 83, 1253–1265.
[2] Weiss, R.A. (2002) IUBMB Life 53, 201–205.
[3] Eckert, D.M. and Kim, P.S. (2001) Annu. Rev. Biochem. 70, 777–810.
[4] Skehel, J.J. and Wiley, D.C. (2000) Annu. Rev. Biochem. 69, 531–569.
[5] Caffrey, M., Cai, M., Kaufman, J., Stahl, S.J., Wingfield, P.T., Covell, D.G., Gronenborn, A.M. and Clore, G.M. (1998) EMBO J. 17, 4572–4584.
[6] Chan, D.C., Fass, D., Berger, J.M. and Kim, P.S. (1997) Cell 89, 263–273.
[7] Tan, K., Liu, J., Wang, J., Shen, S. and Lu, M. (1997) Proc. Natl. Acad. Sci. USA 94, 12303–12308.
[8] Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J. and Wiley, D.C. (1997) Nature 387, 426–430.
[9] Yang, Z.N., Mueser, T.C., Kaufman, J., Stahl, S.J., Wingfield, P.T. and Hyde, C.C. (1999) J. Struct. Biol. 126, 131–144.
[10] Fass, D., Harrison, S.C. and Kim, P.S. (1996) Nat. Struct. Biol. 3, 465–469.
[11] Kobe, B., Center, R.J., Kemp, B.E. and Poumbourios, P. (1999) Proc. Natl. Acad. Sci. USA 96, 4319–4324.
[12] Malashkevich, V.N., Singh, M. and Kim, P.S. (2001) Proc. Natl. Acad. Sci. USA 98, 8502–8506.
[13] Baker, K.A., Dutch, R.E., Lamb, R.A. and Jardetzky, T.S. (1999) Mol. Cell. 3, 309–319.
[14] Zhao, X., Singh, M., Malashkevich, V.N. and Kim, P.S. (2000) Proc. Natl. Acad. Sci. USA 97, 14172–14177.
[15] Malashkevich, V.N., Schneider, B.J., McNally, M.L., Milhollen, M.A., Pang, J.X. and Kim, P.S. (1999) Proc. Natl. Acad. Sci. USA 96, 2662–2667.
[16] Weissenhorn, W., Carfi, A., Lee, K.H., Skehel, J.J. and Wiley, D.C. (1998) Mol. Cell. 2, 605–616.
[17] Blacklow, S.C., Lu, M. and Kim, P.S. (1995) Biochemistry 34, 14955–14962.
[18] Lu, M., Blacklow, S.C. and Kim, P.S. (1995) Nat. Struct. Biol. 2, 1075–1082.
[19] Weissenhorn, W., Calder, L.J., Dessen, A., Laue, T., Skehel, J.J. and Wiley, D.C. (1997) Proc. Natl. Acad. Sci. USA 94, 6065–6069.
[20] Wingfield, P.T., Stahl, S.J., Kaufman, J., Zlotnick, A., Hyde, C.C., Gronenborn, A.M. and Clore, G.M. (1997) Protein Sci. 6, 1653–1660.
[21] Center, R.J., Kobe, B., Wilson, K.A., Teh, T., Howlett, G.J., Kemp, B.E. and Poumbourios, P. (1998) Protein Sci. 7, 1612–1619.
[22] Kong, L.I., Lee, S.W., Kappes, J.C., Parkin, J.S., Decker, D., Hoxie, J.A., Hahn, B.H. and Shaw, G.M. (1988) Science 240, 1525–1529.
[23] Wilson, K.A., Maerz, A.L. and Poumbourios, P. (2001) J. Biol. Chem. 276, 49466–49475.
[24] Rost, B., Fariselli, P. and Casadio, R. (1996) Protein Sci. 5, 1704–1718.
[25] West, J.T., Johnston, P.B., Dubay, S.R. and Hunter, E. (2001) J. Virol. 75, 9601–9612.
[26] Syu, W.J., Lee, W.R., Du, B., Yu, Q.C., Essex, M. and Lee, T.H. (1991) J. Virol. 65, 6349–6352.
[27] Chang, D.K., Cheng, S.F. and Trivedi, V.D. (1999) J. Biol. Chem. 274, 5299–5309.
[28] Li, Y., Han, X. and Tamm, L.K. (2003) Biochemistry 42, 7245–7251.
[29] Fass, D. and Kim, P.S. (1995) Curr. Biol. 5, 1377–1383.
[30] Maerz, A.L., Center, R.J., Kemp, B.E., Kobe, B. and Poumbourios, P. (2000) J. Virol. 74, 6614–6621.
[31] Maerz, A.L., Drummer, H.E., Wilson, K.A. and Poumbourios, P. (2001) J. Virol. 75, 6635–6644.
[32] Poumbourios, P., Maerz, A.L. and Drummer, H.E. (2003) J. Biol. Chem. 278, 42149–42160.
[33] Salzwedel, K., West, J.T. and Hunter, E. (1999) J. Virol. 73, 2469–2480.
[34] Munoz-Barroso, I., Salzwedel, K., Hunter, E. and Blumenthal, R. (1999) J. Virol. 73, 6089–6092.
[35] Park, H.E., Gruenke, J.A. and White, J.M. (2003) Nat. Struct. Biol. 10, 1048–1053.