Published in final edited form as:

Proteins. 2013 February ; 81(2): 316–325. doi:10.1002/prot.24188.

Lethal mutations in the major homology region and their suppressors act by modulating the dimerization of the Rous sarcoma virus capsid protein C-terminal domain

Paula M. Dalessio[†], Rebecca C. Craven[‡], Parvez M. Lokhandwala[‡], and Ira J. Ropson^{†,*}

[†]Department of Biochemistry and Molecular Biology, Penn State University College of Medicine, Hershey, PA 17033

[‡]Department of Microbiology and Immunology, Penn State University College of Medicine, Hershey, PA 17033

Abstract

An infective retrovirus requires a mature capsid shell around the viral replication complex. This shell is formed by about 1500 capsid protein monomers, organized into hexamer and pentamer rings that are linked to each other by the dimerization of the C-terminal domain (CTD). The major homology region (MHR), the most highly conserved protein sequence across retroviral genomes, is part of the CTD. Several mutations in the MHR appear to block infectivity by preventing capsid formation. Suppressor mutations have been identified that are distant in sequence and structure from the MHR and restore capsid formation. The effects of two lethal and two suppressor mutations on the stability and function of the CTD were examined. No correlation with infectivity was found for the stability of the lethal mutations (D155Y-CTD, F167Y-CTD) and suppressor mutations (R185W-CTD, F167Y-CTD). The stabilities of three double mutant proteins (D155Y/R185W-CTD, F167Y/R185W-CTD and F167Y/I190V-CTD) were additive. However, the dimerization affinity of the mutant proteins correlated strongly with biological function. The CTD proteins with lethal mutations did not dimerize, while those with suppressor mutations had greater dimerization affinity than WT-CTD. The suppressor mutations were able to partially correct the dimerization defect caused by the lethal MHR mutations in double mutant proteins. Despite their dramatic effects on dimerization, none of these residues participate directly in the proposed dimerization interface in a mature capsid. These findings suggest that the conserved sequence of the MHR has critical roles in the conformation(s) of the CTD that are required for dimerization and correct capsid maturation.

Keywords

Retroviral capsid protein; protein mutagenesis; protein dimerization; protein stability

INTRODUCTION

CA (capsid) protein is the major structural protein in mature retroviral core structures.¹⁻³ In immature virus particles, as part of the polyprotein Gag, CA provides the intermolecular interactions that organize Gag molecules into the hexameric lattice that is the framework of the immature virus particles. Upon budding from the infected cell, the viral protease cleaves CA and other structural proteins from Gag. A major reorganization of the particle interior

*To whom correspondence should be addressed. 500 University Drive, H171, Hershey, PA 17033. Tel.: 717-531-4064. Fax: 717-531-7072. iropson@psu.edu.

follows, including the formation of the capsid shell around the genome complex to prime the virion for initiation of reverse transcription upon entry into a susceptible cell. Extensive genetic, biochemical and structural studies with several retroviruses have demonstrated that the cleavage and maturation events are accompanied by significant conformational changes in the CA protein and dramatic rearrangements of CA-CA interfaces as the Gag lattice is disassembled and the capsid is formed.¹⁻⁴ Maturation inhibitors that bind CA to block these rearrangements are in development as potential antiretroviral drugs.

The structures of the mature CA proteins from a number of retroviruses are very similar to each other in spite of little sequence conservation.⁵⁻⁸ The CA protein of RSV has been a powerful model system for the study of retroviral capsid assembly.⁹⁻¹² The mature CA protein in Rous sarcoma virus (RSV, 237 residues) consists of two helical domains, the N-terminal domain (NTD, residues 1-147) and C-terminal domain (CTD, residues 151-237) connected by a short linker. Low resolution cryo-electron tomography studies indicate that the RSV capsid shell is organized as a hexameric lattice of CA proteins interrupted with twelve pentameric CA rings which are required to close the protein shell around the virus genome complex.^{13,14} This structural organization is very similar to the recently described HIV capsid model.^{3,15} Intermolecular NTD-NTD and CTD-NTD interactions are required to form these rings, and the rings in turn are linked to each other by dimeric CTD-CTD interactions.^{1-3,13-15}

A highly conserved motif in the CTD, the major homology region (MHR, residues 156-175 in RSV), has critical roles in both Gag assembly and capsid maturation^{9,16-18}. Extensive genetic analyses in RSV have documented differential sensitivities of these two biologic processes to a variety of amino acid substitutions, suggesting that their structural requirements are different. None of the most conserved MHR residues participate directly in any of the CA-CA contacts in the mature capsid, although they appear to control the ability of CA to form capsids during maturation.^{2,14,15} In contrast, recent evidence suggests that some of the most conserved MHR residues may form intermolecular contacts in the Gag lattice.⁴

The MHR motif contains aromatic residues that are buried in the hydrophobic core of the CTD, and conserved polar residues that form a network of hydrogen bonds that likely stabilize the domain. In addition, the MHR sits precisely between the two structural elements that form the CTD-CTD dimer interface.^{5,19-21} This current study draws upon a collection of RSV MHR mutations that cause capsid maturation defects and a set of second-site compensatory mutations to determine whether the known effects on capsid function and virion infectivity can be explained by modulation of protein stability or dimerization capability.

The two virus-lethal MHR mutations examined here (D155Y and F167Y) allow the production of abundant immature virus particles and proper Gag cleavage, but normal core maturation is blocked.^{9,10,22,23} Second site mutations were found that rescued infectivity – R185W rescued the D155Y mutation and I190V rescued F167Y lethality.^{9,10} Suppression of MHR maturation defects is not allele-specific. R185W is a promiscuous suppressor of several lethal MHR mutants while the F167Y replication defect can be corrected by several other secondary mutations including R185W.¹⁰ The control of capsid maturation in mutant particles by these residues was corroborated by use of phosphate anion-driven *in vitro* CA assembly assay. WT-CA assembles via a nucleation-driven process and forms a variety of organized structures that strongly resemble the capsid structures found inside authentic virions.¹² Purified F167Y-CA fails to assemble in this assay and no organized structures are seen by electron microscopy. The D155Y-CA protein retains some assembly capability *in vitro*, but the formation of organized structures is greatly delayed and the structures formed

have altered morphology. Both mutations prevent the formation of stable capsids in virions. The R185W and I190V suppressors promote the nucleation of assembly and formation of capsid-like structures in vitro much more rapidly than WT-CA, can do so at lower phosphate concentration, and in double mutant proteins are able to reverse the assembly defects caused by D155Y and F167Y.¹²

To correlate the effects of these mutations on the biophysical properties of the CTD with the infectivity and in vitro assembly data, the spectra, stability and dimerization activity of purified WT-CTD, two lethal mutations (D155Y-CTD, F167Y-CTD), two suppressors (R185W-CTD, I190V-CTD), and three double mutant proteins (D155Y/R185W-CTD, F167Y/I190V-CTD, F167Y/R185W-CTD) were compared. The findings demonstrate that the residues of the MHR control the ability of CA to form the CTD-CTD dimer interface that is required for capsid formation.

MATERIALS AND METHODS

Protein Source and Purification

Wild-type and D155Y-, D155Y/R185W-, F167Y-, F167Y/I190V-, F167Y/R185W-, I190V-, R185W-CTD sequences were amplified from the previously described proviral pRSV8.eGFP plasmids.¹⁰ Each amplified sequence was then inserted into a pET28-tev plasmid employing *NheI* and *HindIII* restriction sites. This plasmid has a N-terminal 6-His purification tag followed by a tev protease specific cleavage site. The final CTD sequence after tev protease cleavage corresponded to the RSV Prague C sequence (E148 to A237) with five additional amino acids at the N-terminus (GHMAS) from cloning.

The resulting plasmids were transformed into *E. coli* strain BL21-DE3. Most of the his-tagged CTD proteins were expressed using an auto-induction system.²⁴ Transformed *E. coli* were grown in ZYP-5052 medium for approximately 20 hr to ensure that the culture reached saturation and maximize protein production. To produce ¹⁵N-labelled CTD, the bacteria were grown in M9 media supplemented with ¹⁵NH₄Cl. Protein expression in M9 media was induced by addition of 1 mM IPTG after the culture achieved an OD₆₀₀ reading of 1. Cultures were centrifuged at 3100g for 10 min to yield approximately 15 gm of pellet per liter of autoinduction media or 7.5 gm of pellet per liter of M9 media. Pellets were frozen at -80 °C until purification. Each pellet was resuspended in BugBuster© (primary amine-free) (Novagen) containing a complete EDTA-free protease inhibitor cocktail (Roche), Benzonase (Novagen) and Lysozyme (Sigma) according to manufacturer's recommendations and incubated at room temperature for 30 min. Debris from the lysed cells was pelleted by centrifugation at 21000g for 30 min at 4°C. The supernatant was loaded onto a Ni sepharose column (Ni Sepharose High Performance, GE) equilibrated with a buffer containing 25 mM HEPES, 500 mM NaCl, 20 mM imidazole (pH 7.5). The Ni column was washed with buffer containing 25 mM HEPES, 500 mM NaCl, 39 mM imidazole (pH 7.5) to elute non-specifically bound proteins from the column. The His-CTD was eluted from the column with buffer containing 25 mM HEPES, 500 mM NaCl, and 400 mM imidazole (pH 7.5). The His-tag was cleaved from the N-terminus of the CTD by adding a 1:5 mass ratio of purified Tev protease²⁵ to the purified protein and incubated for 4 hr at room temperature. The solution was dialyzed overnight at 4°C against a buffer containing 25 mM Hepes, 200 mM NaCl (pH 7.5). The protein mix (cleaved CTD, His-tagged peptide, and His-tagged Tev) was then loaded onto a Ni column pre-equilibrated with buffer containing 25 mM HEPES, 500 mM NaCl, 20 mM imidazole (pH 7.5) and washed with the same buffer. The purified CTD, found in the void volume, was concentrated further using ICON© concentrators (5K MWCO, Pierce) and then dialyzed against a buffer containing 10 mM NaPO₄, 0.1 mM EDTA, 5 mM NaN₃, (pH 7.5). A yield of 40 – 80 mg of purified CTD per liter of media was obtained. Purity was monitored by SDS-Page stained with Coomassie blue. In most

preparations, no further purification was necessary to obtain >98% pure protein. In a few cases, an additional size exclusion chromatography step (Superdex 75) equilibrated with 20 mM Tris-HCl, 150 mM NaCl, and 0.1 mM EDTA, (pH 7.5) was necessary to obtain >98% pure protein. The extinction coefficients for each protein were estimated from the content of tyrosine, phenylalanine and tryptophan in each sequence (WT, 6990 M⁻¹cm⁻¹; D155Y, 8480 M⁻¹cm⁻¹; F167Y, 8480 M⁻¹cm⁻¹; R185W, 12490 M⁻¹cm⁻¹; I190V, 6990 M⁻¹cm⁻¹; D155Y/R185W, 13,980 M⁻¹cm⁻¹; F167Y/I190V, 8480 M⁻¹cm⁻¹; F167Y/R185W, 13,980 M⁻¹cm⁻¹).²⁶

Reagents

Guanidine hydrochloride (Gdn-HCl) stock solutions (approximately 8 M, ultrapure, MP Biomedical) were prepared in H₂O and stored at -20°C. On the day of an experiment, a working solution of 6 M Gdn-HCl was prepared by adding the buffer components to a freshly thawed Gdn-HCl stock solution. The buffer at pH 7.0 contained 25 mM NaPO₄, 75 mM NaCl, and 0.1 mM EDTA. The buffer at pH 3.70 contained 20 mM succinate, 100 mM KCl, and 0.1 mM EDTA. The pH of the Gdn-HCl solutions were adjusted to pH 7.0 and pH 3.7, respectively. Actual denaturant concentrations were determined by refractive index measurements with a Milton Roy Abbe-3 refractometer at 25 °C.²⁷ All buffers were filtered through 0.2 µm Whatman nylon membranes. All other chemicals were reagent grade.

Spectral Measurements

Fluorescence and CD spectra were collected for each protein (0.015 m M) at 0 and 6 M GdnHCl in 25 mM NaPO₄, 75 mM NaCl, and 0.1 mM EDTA at 25 °C. All spectra were corrected for buffer contributions. For fluorescence, a PTI QuantaMaster luminescence spectrometer was used. The excitation wavelength was at 290 nm (2 nm bandpass) and emission spectra were collected from 300 to 400 nm (4 nm bandpass) in a 0.3 cm path length cell. A Jasco J-710 spectropolarimeter was used to examine the far-UV portion of the CD spectrum (typically 250 – 200 nm) using a thermostatted 0.1 mm cell at 25 °C. The CDPro²⁸ and K2D3²⁹ programs were used to estimate secondary structure content for the WT and mutant proteins from the CD spectra.

Equilibrium Unfolding Studies

Unfolding transitions as a function of denaturant concentration were monitored by circular dichroism (CD). A Jasco J-710 spectropolarimeter was employed to follow changes in secondary structure in the wavelength range of 225-212 nm using a 0.1 mm cell at 25 °C. Samples ranging in denaturant concentrations from 0 to 6 M Gdn-HCl were prepared using a Hamilton Microlab titrator immediately before an experiment.³⁰ The final protein concentration of was typically 150 µg/ml (0.015 mM). The data were corrected for the background signal of the buffer and Gdn-HCl solutions. For comparison of denaturation curves, the data were normalized from 0 to 1, where 0 and 1 were assigned to the lowest and highest observed intensity, respectively. Replicate experiments for each protein at each pH overlaid well. No significant differences in the unfolding profile were observed for ¹⁵N-labeled and unlabeled CTD proteins (data not shown).

Computer Fitting of Equilibrium Unfolding Data

Non-linear least-square fit to the equilibrium unfolding data were generated by using KaleidaGraph (Synergy Software) in conjunction with the following equation;³¹

$$X_{[D]} = \frac{\{X_N + m_N [D]\} + \{X_U + m_U [D] \exp [(-\Delta G_{H_2O} + m_G [D]) / RT]\}}{\{1 + \exp [(-\Delta G_{H_2O} + m_G [D]) / RT]\}}$$

where $X_{[D]}$ is the value of the spectroscopic property at some denaturant concentration, X_N and X_U are the values for the spectroscopic property linearly extrapolated to $[D] = 0$ for the native and unfolded forms of the protein, respectively, m_N and m_U are the slopes for the dependence of the spectroscopic properties X_N and X_U on denaturant concentration, respectively, ΔG_{H_2O} is the apparent free energy difference between the folded and unfolded forms of the protein linearly extrapolated to $[D] = 0$, m_G is the slope describing the dependence of ΔG_{H_2O} on $[D]$, assuming a linear relationship between the log of the equilibrium constant and denaturant constant, R is the gas constant, and T is the temperature. The midpoint of the transition was determined by substituting $\Delta G_{H_2O}/\text{midpoint}$ for m_G in the equation shown above. The unfolding equilibrium at a number of protein concentrations and dissociation constants was simulated in Mathematica, using the equations that describe the linked equilibria (Dimer $\rightleftharpoons$ Folded Monomer $\rightleftharpoons$ Unfolded monomer).³² The values for ΔG (4.72 kcalmol⁻¹) and m_G (-1.82 kcalmol⁻¹M⁻¹) for the unfolding of WT-CTD at pH 3.7 and 25 °C were used to simulate the potential effects of dimerization at various protein concentrations on the unfolding equilibrium.

Isothermal Titration Calorimetry

ITC experiments were performed with a VP-ITC microcalorimeter (MicroCal) at 10° C in a manner similar to that described by Bailey et al.¹⁹ Briefly, each CTD protein was exhaustively dialyzed against 20 mM succinic acid, 100 mM NaCl, and 5 mM NaN₃ at pH 3.7. Protein and buffer solutions were degassed, and the protein solution (0.35 to 0.5 mM) was introduced into the injection syringe. 25 injections were performed for each titration (1 – 4 µl, 10 – 5µl, 14 – 10 µl). Origin (MicroCal) was used to integrate the volume of heat generated per injection and to fit the data to a monomer – dimer equilibrium model for the dissociation process.³³ The first data point was discarded from the data analysis. Duplicate or greater replicate titrations were performed for each protein.

Surface exposure of the WT residues mutated in this study were made using getArea.³⁴ The percent exposure of each sidechain in the 3g1g monomer and the 3g21 dimer structures¹⁹ were calculated relative to the solvent exposure of that sidechain in a Gly-X-Gly sequence.

RESULTS

The spectral properties and stabilities of WT and mutant CTDs

The fluorescence and CD spectra of the WT and mutant proteins were collected to assess the effects of the mutations on the tertiary and secondary structure of the domain. The fluorescence spectra of all of the proteins are shown in Figure 2A. WT-CTD has a single tryptophan (W153), near the N terminus of the structure (Figure 1) and close in sequence to the lethal D155Y mutation. However, the sidechains of D155 and W153 are not oriented towards each other, and are unlikely to interact directly with each other after mutation. The W153 sidechain projects towards helix 2, and is mostly buried in a hydrophobic pocket in the monomeric CTD structure (3g1g.pdb).¹⁹ Only minor changes in fluorescence intensity and wavelength of maximal fluorescence were observed for most mutations, suggesting that little if any tertiary structure changes had occurred in the vicinity of W153. For the three mutant proteins encoding an additional tryptophan (R185W-CTD, F167Y/R185W-CTD, D155Y/R185W-CTD), large increases in fluorescence intensity were observed compared to

WT, but the wavelength of maximal fluorescence was unchanged, suggesting that the mutated sidechain has limited solvent exposure in these proteins as well.

The CD spectra were typical for helical proteins, showing the diagnostic minima at 207 and 223 nm (Figure 2B). However, there were significant changes in intensity that correlated with changes in aromatic residues. The two proteins containing the F167Y mutation (F167Y-CTD, F167Y/I190V-CTD) had slightly less intensity at 223 nm than WT-CTD, while those proteins containing an additional tryptophan (R185W-CTD, F167Y/R185W-CTD, D155Y/R185W-CTD) showed much greater intensity than WT-CTD at 223 nm. Two programs (CDPro²⁸ and K2D3²⁹) were used to estimate the secondary structure content of the WT-CTD and mutant-CTD proteins from their CD spectra. Both programs predicted a lower alpha helical content (48%) for the WT-CTD spectra than expected from crystallography (58%¹⁹). All of the mutant spectra have predicted helical structure similar to that of the WT protein, with those spectra having higher signal at 223 nm (proteins with the R185W mutation) having 2-4% higher helical content, whereas the spectra having lower intensity at 223 nm (proteins with the F167Y mutation) had about 2% lower helical content. Extending the spectral range to 190 nm for the WT-CTD, F167Y-CTD and R185W-CTD proteins did not change the estimates of secondary structure for those proteins significantly. Although it appears unlikely that any large changes in secondary structure were made by any of the mutations, small changes in secondary structure may still have occurred. Alternatively, aromatic residues can contribute to the far UV CD spectra of a protein,³⁵ and these intensity changes at 223 nm may be due to new or altered aromatic side chain interactions in these proteins.

The unfolding of the CTD domain is shown in Figure 3. The reaction appeared to be completely reversible, so long as the incubation time in the unfolded state was limited to less than two hours (data not shown). WT-CTD has 10 prolines, all of which are *trans* in the native state structures. Prolonged incubation in the unfolded state is likely to cause extensive proline isomerization,³⁶ and appears to limit the reversibility of folding for this domain after long incubations in the unfolded state or after thermal denaturation.

The stabilities of WT-CTD, the lethal F167Y-CTD mutation, the R185W-CTD suppressor, and the F167Y/R185W-CTD double mutant proteins are compared in Figure 3A. The F167Y mutation destabilized the domain, while the R185W had the opposite effect. Once again, the stability of the double mutant protein was nearly identical to that expected from the sum of the stabilities of each mutation alone (calculated ΔG of 0.58 ± 0.33 kcal/mol compared to the measured change of 0.66 ± 0.26 kcal/mol).

The stability of WT-CTD, the lethal D155Y-CTD mutation, the R185W-CTD suppressor, and the double mutant D155Y/R185W-CTD are compared in Figure 3B. The D155Y mutation had very little effect on CTD stability whereas the R185W mutation stabilized it considerably. The stability of the D155Y/R185W double mutant was nearly additive for the two single site mutations (calculated $\Delta\Delta G$ of 1.66 ± 0.34 kcal/mol compared to the measured change of 1.81 ± 0.22 kcal/mol), well within the error of measurement (Table I).

The stability of WT-CTD, the lethal F167Y-CTD mutation, the I190V-CTD suppressor of this mutation, and the double mutant F167Y/I190V-CTD are compared in Figure 3C. Both F167Y-CTD and I190V-CTD were destabilized compared to WT-CTD (Table I), and the stability of the double mutant F167Y/I190V-CTD was similar to the value expected for the sum of the effects of each individual mutation (calculated ΔG of -0.91 ± 0.30 kcal/mol compared to the measured change of -1.36 ± 0.19 kcal/mol).

One potential complication in the measurement of the stability of these proteins is dimerization, especially at pH 3.7 and 25 °C, as described below. Unlike HIV-CTD, there is

no measurable dimerization of RSV-CTD at neutral pH.¹⁹ At pH 3.7, the dissociation constant at 20 °C was about 0.25 mM.¹⁹ The absence of dimerization at neutral pH and the dissociation constants at 20 and 25 °C at pH 3.7 have been confirmed in our studies (data not shown). In order to determine if dimerization will affect unfolding at 0.015 mM protein concentration and 25 °C, the effects of dimerization on measured stability were simulated given the thermodynamic parameters for monomer unfolding and the measured binding affinities at pH 3.7 and equations³² describing the linked equilibria (Figure S1). Further, we measured the stability of the WT-CTD domain at a much higher protein concentration (0.2 mM, 2 mgs/ml) at pH 3.7 (Figure 4A). As expected from the simulations, there was no significant difference in the stability of WT-CTD at either concentration of protein. The unfolding of WT-CTD is accurately described by a simple two-state model for the unfolding process at pH 3.7 (Native monomer $\rightleftharpoons$ Unfolded monomer). Dimerization may affect the measurement of the stability of the tight binding suppressor mutations (R185W-CTD and I190V-CTD) at pH 3.7, but the simulations suggest that the effects would be minimal at 0.015 mM protein concentration (Figure S1). All of the fits reported here are to a simple two-state model for the unfolding transition, with no contribution to the equilibrium from dimerization. Although a slightly lower midpoint for denaturation was observed at pH 3.7 compared to pH 7.0 for the four proteins studied (WT-CTD, F167Y-CTD, R185W-CTD, F167Y/R185W-CTD, Figure 4, Table II), the relative effects of the mutations on stability were nearly identical to those observed at pH 7.0. Further, the CD spectra at pH 3.7 were identical to those collected at pH 7.0, suggesting that no significant conformation or stability changes occurred for the monomeric form of any of these proteins at lower pH values.

Dimerization of WT and Mutant CTDs

To assess the ability of mutant CTDs to dimerize, the heat of dissociation of dimers was examined by isothermal titration calorimetry at 10 °C and pH 3.7.¹⁹ These heats were generated by dilution of 0.5 mM protein equilibrated at pH 3.7 into buffer. Figure 5 shows the heat generated during a typical experiment for WT-CTD under these conditions. Figure 6 shows a comparison of the integrated heats of dissociation for WT-CTD, F167Y-CTD, R185W-CTD and F167Y/R185W-CTD. The equilibrium dissociation constant (K_d) for WT-CTD (0.11 mM) is similar to that reported previously for a slightly different sequence of the same domain.¹⁹ Neither the F167Y-CTD (Figure 6) nor the D155Y-CTD proteins showed any significant curvature, suggesting that no detectable dimer was present at 0.5 mM concentration. Under these conditions, the weakest K_d that can be observed is about 1.5 mM. As such, both lethal mutants appear to have at least one order of magnitude weaker dimerization affinity than the WT-protein. The suppressor mutations alone (R185W-CTD and I190V-CTD; Table I and Figure 6) associated more strongly than the WT-protein. The double mutant proteins (D155Y/R185W-CTD, F167Y/R185W-CTD and F167Y/I190V; Table I and Figure 6) associate more weakly than the WT-protein, but dimerization still occurs when the suppressor mutations are combined with the lethal mutations.

DISCUSSION

Structure and Stability of the CTD

No dramatic changes in the fluorescence and CD spectra were observed for either the D155Y-CTD or the F167Y-CTD lethal mutant proteins compared to the WT-CTD, suggesting that any changes in the overall structure were minimal. Likewise the I190V suppressor, a very conservative substitution, caused little detectable perturbation. Of these, only the F167Y mutation resulted in a distinct decrease in domain stability. This decrease in stability correlates well with the weak infectivity of the F167Y virus at 37 °C and the complete loss of infectivity at 42 °C.¹⁰ Although this mutation is conservative, it is in the

center of the small hydrophobic core of the domain, and could easily have a destabilizing effect.

The R185W suppressor mutation, however, appears to cause greater structural changes. The wavelength of maximal fluorescence emission (338 nm, Figure 2A) of the mutant protein suggests that the additional tryptophan is more buried in a hydrophobic environment than would be expected from the surface orientation of the R185 side chain. The R185 side chain participates in an intra-domain salt bridge with E176,¹⁹ and this interaction would be lost when this residue is replaced with tryptophan. As such, the observed stabilization of the domain by the R185W substitution was unexpected, and some conformation changes must have occurred that both increases the stability of the domain and strengthens the dimerization interface. This stabilization explains the previously observed biphasic equilibrium unfolding profile for the R185W mutation in the context of full length RSV-CA.¹² The far-UV CD spectra of all of the proteins containing the R185W mutation was perturbed (Figure 2B), especially at 223 nm, suggesting the possibility of a slight increase in the α -helical content in the R185W single and double mutant CTDs. Stabilization of the dimerization helix by the R185W mutation is consistent with the overall stabilization of the domain by this mutation, the promiscuous suppression of many lethal mutations by the R185W mutation *in vivo*,¹⁰ the biophysical³⁷ and modeling³⁸ studies of the dimerization of HIV-CTD and the structural studies of Mason-Pfizer monkey virus particles.⁴ Although the magnitude of the effect is small, and we cannot be confident that it has biological relevance, it does suggest a possible mechanism of action for the R185W mutation.

The double mutant proteins (D155Y/R185W-CTD, F167Y/I190V-CTD, and F167Y/R185W-CTD) had stabilities that were nearly exactly additive for each of the single site mutations alone (Tables I and II). In the F167Y/R185W-CTD double mutant, the R185W suppressor completely overcame the destabilizing effect of the lethal mutation, resulting in a protein that was more stable than WT-CTD. However the relevance of these observations to the biological effects of these mutations *in vivo* (and by extension to the function of the MHR motif) is limited given that the I190V mutation, which is also a robust suppressor of F167Y, slightly destabilizes the protein. As such, the mechanism of action of these suppressors may differ, and there is no simple correlation between the effects of the mutations on domain stability and their effects on infectivity.

Dimerization of the CTD

Dimerization between the second α -helices of retroviral CTDs is a critical structural feature in all retroviral systems examined. This interaction is important to the formation of Gag hexamers in the immature lattice and between CA subunits of neighboring hexamers and pentamers in the mature capsid shell.¹⁻³ However, there is remarkable plasticity in the intermolecular dimer interface.³⁷ The crossing angle of these two helices changes dramatically between the immature lattice and the mature capsid shell for both HIV¹⁻³ and RSV.^{13,14,19} A recent study on an 8 Å resolution structure of Gag protein of the related Mason-Pfizer monkey retrovirus concludes that the two CTDs rotate a remarkable 119 degrees about this interface during maturation.⁴ Proteolysis initiates the maturation process, but the structure factors that direct this reorganization have been very poorly defined.

Although CTD dimerization is a feature of all orthoretroviral capsids, the isolated CA proteins typically dimerize only reluctantly.³ The HIV protein is somewhat exceptional in that both HIV-CA and HIV-CTD dimerize at neutral pH although with a rather weak affinity. We have detected assembly-capable dimers of full-length CA that are unstable in purified form.¹¹ Assembly of capsid structures is electrostatically controlled in both HIV^{15,39,40} and RSV¹⁰⁻¹². Full length RSV protein can be triggered to assemble into capsid-like structures of apparently identical organization either by exposure to low pH, which

provides for protonation of one of two aspartate residues on the dimerization helix,^{19,41} or by exposure to high anion concentration at neutral pH which presumably acts by shielding the large positive surface charge on the NTD.^{11,12,14} Crystallization of the RSV CTD in low pH yielded the dimeric structure that is consistent with the recently developed models for the CA lattice of HIV¹⁵ and RSV¹³ capsids. The X-ray crystal structures of HIV CTD dimer^{5,21,42} and the low pH CTD dimer for RSV¹⁹ are quite similar to each other and are accepted in the field as the best representations of the dimer interface in mature capsids.^{2,3,15} The factor(s) that trigger assembly of CA in maturing virions is unknown although it is likely that other viral components, such as RNA, serve this role. Nevertheless, low pH appears to be a reasonable surrogate for *in vitro* studies.

The effects of these MHR and suppressor mutations on the dimerization of the CTD domain at low pH fully explain the observed infectivity data *in vivo*^{9,10,22} (Table I) and the assembly of purified full-length CA protein into capsid-like structures *in vitro*.¹¹ The suppressor mutations increase the dimerization affinity of the CTD, and increase the rate of formation of capsid-like structures from purified protein *in vitro*, suggesting that the rate of dimer formation is one of the rate limiting steps in capsid formation.¹¹

None of the mutations described in this study participate directly in the dimerization interface. However, the small size of the domain (90 residues) places all of the mutation sites in close proximity to one or more residues that do participate. D155 is not a highly conserved residue, and is located at the boundary of the MHR motif and the upstream 3¹⁰ helix that participates in the mature dimer interface. D155 itself does not appear to have a significant role in the interface and is largely solvent exposed (70%) in the dimer, and the mechanism by which dimerization is prevented is unclear. It is possible that a relatively minor conformation change might preclude dimerization, perhaps by flipping the orientation of the hydrophobic tyrosine side chain into the dimer binding interface. The highly conserved F167 is buried in the hydrophobic core of the CTD where it could influence the CTD-CTD interface by inducing conformation changes without being directly involved. The mechanism by which the suppressor mutations increase dimerization affinity is also unclear. Both the R185W and I190V suppressors are located on the second α -helix that is critical to dimerization, but their side chains are on the opposite side of the helix from the dimer interface. The promiscuous suppression behavior of the R185W mutation is particularly interesting, since the side chain of the R185 residue is oriented towards solvent away from the dimer interface. The changes in the CD and fluorescence spectra caused by the R185W mutation are larger than those for the other mutations, suggesting that significant changes in the structure of the domain could have occurred, perhaps by reorienting the mutant tryptophan side chain towards the dimer interface or by stabilizing the dimerization helix as described above.

Complex behavior for the dimerization of HIV-CTD has also been observed. Multiple mutations in HIV-CTD in this vicinity had non-additive effects on dimerization, suggesting that this interface has complicated energetics.³⁹ A computational study of four other mutations in HIV-CTD suggest that the conformation of the MHR is critical to the correct function of the dimerization interface, even though those residues do not directly participate in the interface.^{37,38} These RSV studies are a direct test of that prediction and emphasize the utility of biologically derived suppressors in examining capsid interface surfaces.

The work described here argues that a critical function of the conserved MHR region of CA controls the ability of the CTD to form the mature dimer interface found in the capsids of mature infective virions. It is a critical key to understanding the regulation of maturation and relevant to development of retroviral maturation inhibitors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Joerg Rosgen for his assistance in modeling the linked dimerization and unfolding equilibrium in Mathematica. This research was supported by NIH grant AI079694 (R.C.C. and I.J.R.) and CA100322 (R.C.C.) and funds from the Pennsylvania Department of Health to R.C.C.

Abbreviations

RSV	Rous sarcoma virus
HIV	Human immunodeficiency virus
CA	virus capsid protein
CTD	C-terminal domain of the virus capsid protein
CD	circular dichroism
ΔG_{H_2O}	free energy of stabilization in the absence of denaturant
MHR	major homology region
m_G	dependence of ΔG on denaturant concentration
Kd	equilibrium dissociation constant

REFERENCES

1. Ganser-Pornillos BK, Yeager M, Sundquist WI. The structural biology of HIV assembly. *Curr Opin Struct Biol.* 2008; 18(2):203–217.
2. Mateu MG. The capsid protein of human immunodeficiency virus: intersubunit interactions during virus assembly. *FEBS J.* 2009; 276(21):6098–6109. [PubMed: 19825044]
3. Ganser-Pornillos BK, Yeager M, Pornillos O. Assembly and architecture of HIV. *Adv Exp Med Biol.* 2012; 726:441–465. [PubMed: 22297526]
4. Bharat TA, Davey NE, Ulbrich P, Riches JD, de Marco A, Rumlova M, Sachse C, Ruml T, Briggs JA. Structure of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy. *Nature.* 2012; 487(7407):385–389. [PubMed: 22722831]
5. Byeon IJ, Meng X, Jung J, Zhao G, Yang R, Ahn J, Shi J, Concel J, Aiken C, Zhang P, Gronenborn AM. Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell.* 2009; 139(4):780–790. [PubMed: 19914170]
6. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M. X-ray structures of the hexameric building block of the HIV capsid. *Cell.* 2009; 137(7):1282–1292. [PubMed: 19523676]
7. Kingston RL, Fitzon-Ostendorf T, Eisenmesser EZ, Schatz GW, Vogt VM, Post CB, Rossmann MG. Structure and self-association of the Rous sarcoma virus capsid protein. *Struct Fold Des.* 2000; 8(6):617–628.
8. Khorasanizadeh S, Campos-Olivas R, Summers MF. Solution structure of the capsid protein from the human T-cell leukemia virus type-I. *Journal of Molecular Biology.* 1999; 291(2):491–505. [PubMed: 10438634]
9. Bowzard JB, Wills JW, Craven RC. Second-site suppressors of Rous sarcoma virus CA mutations: evidence for interdomain interactions. *J Virol.* 2001; 75(15):6850–6856. [PubMed: 11435564]
10. Lokhandwala PM, Nguyen TL, Bowzard JB, Craven RC. Cooperative role of the MHR and the CA dimerization helix in the maturation of the functional retrovirus capsid. *Virology.* 2008; 376(1):191–198. [PubMed: 18433823]

11. Purdy JG, Flanagan JM, Ropson IJ, Craven RC. Retroviral capsid assembly: a role for the CA dimer in initiation. *J Mol Biol.* 2009; 389(2):438–451. [PubMed: 19361521]
12. Purdy JG, Flanagan JM, Ropson IJ, Rennoll-Bankert KE, Craven RC. Critical role of conserved hydrophobic residues within the major homology region in mature retroviral capsid assembly. *J Virol.* 2008; 82(12):5951–5961. [PubMed: 18400856]
13. Heymann JB, Butan C, Winkler DC, Craven RC, Steven AC. Irregular and semi-regular polyhedral models for Rous sarcoma virus cores. *Computational and Mathematical Methods in Medicine.* 2008; 9:197–210. [PubMed: 19122884]
14. Cardone G, Purdy JG, Cheng N, Craven RC, Steven AC. Visualization of a missing link in retrovirus capsid assembly. *Nature.* 2009; 457(7230):694–698. [PubMed: 19194444]
15. Pornillos O, Ganser-Pornillos BK, Yeager M. Atomic-level modelling of the HIV capsid. *Nature.* 2011; 469(7330):424–427. [PubMed: 21248851]
16. Mammano F, Ohagen A, Hoglund S, Gottlinger HG. Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J Virol.* 1994; 68:4927–4936. [PubMed: 8035491]
17. Strambio de Castillia C, Hunter E. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. *J Virol.* 1992; 66
18. Willems L, Kerkhofs P, Attenelle L, Burny A, Portetelle D, Kettmann R. The major homology region of bovine leukaemia virus p24gag is required for virus infectivity in vivo. *Journal of General Virology.* 1997; 78(Pt 3):637–640. [PubMed: 9049415]
19. Bailey GD, Hyun JK, Mitra AK, Kingston RL. Proton-linked dimerization of a retroviral capsid protein initiates capsid assembly. *Structure.* 2009; 17(5):737–748. [PubMed: 19446529]
20. Bailey GD, Hyun JK, Mitra AK, Kingston RL. A structural model for the generation of continuous curvature on the surface of a retroviral capsid. *Journal of molecular biology.* 2012; 417(3):212–223. [PubMed: 22306463]
21. Worthylake DK, Wang H, Yoo S, Sundquist WI, Hill CP. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr D Biol Crystallogr.* 1999; 55(Pt 1):85–92. [PubMed: 10089398]
22. Cairns TM, Craven RC. Viral DNA synthesis defects in assembly-competent Rous sarcoma virus CA mutants. *J Virol.* 2001; 75(1):242–250. [PubMed: 11119594]
23. Butan C, Winkler DC, Heymann JB, Craven RC, Steven AC. RSV capsid polymorphism correlates with polymerization efficiency and envelope glycoprotein content: implications that nucleation controls morphogenesis. *Journal of Molecular Biology.* 2008; 376(4):1168–1181. [PubMed: 18206161]
24. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein ExprPurif.* 2005; 41(1):207–234.
25. Kapust RB, Tozser J, Fox JD, Anderson DE, Cherry S, Copeland TD, Waugh DS. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* 2001; 14(12):993–1000. [PubMed: 11809930]
26. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 1995; 4(11):2411–2423. [PubMed: 8563639]
27. Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* 1986; 131:266–280. [PubMed: 3773761]
28. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem.* 2000; 287(2):252–260. [PubMed: 11112271]
29. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins.* 2012; 80:374–381.
30. Gualfetti PJ, Bilsel O, Matthews CR. The progressive development of structure and stability during the equilibrium folding of the alpha subunit of tryptophan synthase from *Escherichia coli*. *Protein Sci.* 1999; 8(8):1623–1635. [PubMed: 10452606]
31. Santoro MM, Bolen DW. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry.* 1988; 27(21):8063–8068. [PubMed: 3233195]

32. Maity H, Mossing MC, Eftink MR. Equilibrium unfolding of dimeric and engineered monomeric forms of lambda Cro (F58W) repressor and the effect of added salts: evidence for the formation of folded monomer induced by sodium perchlorate. *Arch Biochem Biophys*. 2005; 434(1):93–107. [PubMed: 15629113]
33. Burrows SD, Doyle ML, Murphy KP, Franklin SG, White JR, Brooks I, McNulty DE, Scott MO, Knutson JR, Porter D, et al. Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry*. 1994; 33(43):12741–12745. [PubMed: 7947677]
34. Fraczekiewicz R, Braun W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem*. 1998; 19(3):319–333.
35. Woody RW. Circular dichroism. *Methods Enzymol*. 1995; 246:34–71. [PubMed: 7538625]
36. Reimer U, Scherer G, Drewello M, Kruber S, Schutkowski M, Fischer G. Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *Journal of Molecular Biology*. 1998; 279(2):449–460. [PubMed: 9642049]
37. Mateu MG. Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus-1 capsid protein. *Journal of molecular biology*. 2002; 318(2):519–531. [PubMed: 12051856]
38. Yu X, Wang Q, Yang JC, Buch I, Tsai CJ, Ma B, Cheng SZ, Nussinov R, Zheng J. Mutational analysis and allosteric effects in the HIV-1 capsid protein carboxyl-terminal dimerization domain. *Biomacromolecules*. 2009; 10(2):390–399. [PubMed: 19199580]
39. del Alamo M, Mateu MG. Electrostatic repulsion, compensatory mutations, and long-range non-additive effects at the dimerization interface of the HIV capsid protein. *J Mol Biol*. 2005; 345(4): 893–906.
40. del Alamo M, Neira JL, Mateu MG. Thermodynamic dissection of a low affinity protein-protein interface involved in human immunodeficiency virus assembly. *J Biol Chem*. 2003; 278(30): 27923–27929.
41. Hyun JK, Radjainia M, Kingston RL, Mitra AK. Proton-driven assembly of the Rous Sarcoma virus capsid protein results in the formation of icosahedral particles. *J Biol Chem*. 2010; 285(20): 15056–15064. [PubMed: 20228062]
42. Gamble TR, Yoo S, Vajdos FF, von Schwedler UK, Worthylake DK, Wang H, McCutcheon JP, Sundquist WI, Hill CP. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science*. 1997; 278(5339):849–853. [PubMed: 9346481]

Figure 1.

Structure and sequence of the Rous sarcoma virus CA C-terminal domain construct. The structure shows the termini and mutation sites in the dimeric low pH biological unit (3g1g.pdb1, G151-K227). The termini and mutation sites are shown on each monomer. The sequence numbering is from the full-length CA-protein and shows the location of the interdomain linker (IDL), the major homology region (MHR), the N and C terminus of the defined structural region, and the helices. Mutation sites are italicized.

Figure 2.

A. Fluorescence spectra of native WT- and mutant CTD proteins at pH 7. B. CD spectra of native WT- and mutant CTD proteins at pH 7.

Figure 3.

A. The unfolding of WT-CTD, F167Y-CTD, R185W-CTD, and F167Y/R185W-CTD at pH 7. B. The unfolding of D155Y-CTD, R185W-CTD, and D155Y/R185W-CTD at pH 7. C. The unfolding of F167Y-CTD, I190V-CTD, and F167Y/I190V-CTD at pH 7. The lines through the data are the fits to a two state model for the unfolding transition. The two-state fit for WT-CTD is shown as a gray line in panels B and C for comparison.

Figure 4.

A. The unfolding of 0.2 and 0.015 mM WT-CTD at pH 3.7. The line through the data is the fit to a two state model for the unfolding transition at 0.015 mM protein concentration. B. The unfolding of D155Y-CTD, R185W-CTD, and D155Y/R185W-CTD at pH 3.7. The lines through the data are the fits to a two state model for the unfolding transition. The two-state fit for WT-CTD is shown as a gray line for comparison.

Figure 5.

Typical calorimetry data, resulting from the sequential dilution of 500 μM WT-CTD into buffer at 10 $^{\circ}\text{C}$ and pH 3.7. An initial injection of 4 μl was followed by 10 injections of 5 μl and 14 injections of 10 μl .

Figure 6. Integrated isothermal titration data for the injection of WT-CTD and mutant proteins at 10 °C and pH 3.7. The lines are fits of the titration data to a monomer-dimer equilibrium.

Parameters for the Unfolding at pH 7, Dimerization at pH 3.7 and Infectivity of WT and Mutant CTD Sequences

Table I

Protein	ΔG_{H_2O} (kcalmol ⁻¹)	mG (kcalmol ⁻¹ M ⁻¹)	Midpoint (M)	$\Delta(\Delta G_{H_2O})$ (kcalmol ⁻¹)	K _d (mM)	Relative Infectivity
WT-CTD	4.72 ± 0.16	-1.75 ± 0.06	2.70 ± 0.02	-	0.11 ± 0.03	1.00
¹⁵ N-D155Y-CTD	5.06 ± 0.16	-1.90 ± 0.06	2.66 ± 0.01	0.34 ± 0.23	>1.5	< 0.05
F167Y-CTD	3.98 ± 0.13	-1.92 ± 0.06	2.08 ± 0.01	-0.74 ± 0.21	>1.5	0.08
R185W-CTD	6.04 ± 0.19	-1.70 ± 0.05	3.55 ± 0.01	1.32 ± 0.23	0.02 ± 0.003	0.81
I190V-CTD	4.55 ± 0.13	-1.81 ± 0.05	2.52 ± 0.01	-0.17 ± 0.21	0.04 ± 0.01	0.74
¹⁵ N-D155Y/R185W-CTD	6.53 ± 0.15	-1.86 ± 0.04	3.51 ± 0.01	1.81 ± 0.22	0.99 ± 0.28	0.54
F167Y/I190V-CTD	3.36 ± 0.10	-1.82 ± 0.06	1.84 ± 0.01	-1.36 ± 0.19	0.55 ± 0.19	0.89
F167Y/R185W-CTD	5.38 ± 0.21	-1.86 ± 0.07	2.90 ± 0.01	0.66 ± 0.26	0.54 ± 0.23	0.98

Proteins were unfolded with Gdn-HCl and fit to a two state model for the unfolding process. $\Delta(\Delta G_{H_2O})$ is the difference in stability between the mutant and WT protein at 0 M denaturant at pH 7. A positive $\Delta(\Delta G_{H_2O})$ indicates stabilization by the mutant. There is no detectable dimer formation at neutral pH for any of the proteins. The K_d is the dimer dissociation constant at pH 3.7. Relative infectivity of a mutant viral sequence compared to the WT sequence in cell culture.⁸

Table II

Parameters for the Equilibrium Unfolding of WT and Mutant CTD Proteins at pH 3.7

Protein	ΔG_{H_2O} (kcalmol ⁻¹)	mG (kcalmol ⁻¹ M ⁻¹)	Midpoint (M)	$\Delta(\Delta G_{H_2O})$ (kcalmol ⁻¹)
WT-CTD (0.015 mM)	4.31 ± 0.14	-1.82 ± 0.06	2.37 ± 0.02	
WT-CTD (0.20 mM)	4.44 ± 0.11	-1.83 ± 0.07	2.43 ± 0.02	
F167Y-CTD	3.30 ± 0.13	-1.79 ± 0.05	1.84 ± 0.02	-1.01 ± 0.19
R185W-CTD	5.95 ± 0.19	-1.78 ± 0.06	3.35 ± 0.01	1.64 ± 0.24
F167Y/R185W-CTD	5.40 ± 0.23	-1.95 ± 0.08	2.77 ± 0.02	1.09 ± 0.27

Proteins were unfolded with Gdn-HCl and fit to a two state model for the unfolding process. $\Delta(\Delta G_{H_2O})$ is the difference in stability between the mutant and WT protein at 0 M denaturant. A positive $\Delta(\Delta G_{H_2O})$ indicates stabilization by the mutant.