with experiment. This agreement suggests that our procedure accurately evaluates λ and V for these systems. It is noteworthy that, for the strongly adiabatic systems studied here, classical theory that ignores tunneling reproduces the observed rate constant, even though $h\nu_v$ modes far larger than $2RT$ are involved. We suggest that this success means that tunneling effects have been effectively incorporated into the calculated ground-state adiabatic surface by our band-fitting procedure. To our knowledge, there have been no prior comparisons of calculated rate constants based on optically determined ET parameters with experimental ones for ET processes that occur in the adiabatic rate regime. Our results verify that Hush's method for extracting V is quantitatively correct in the absence of intensity borrowing problems (*10*). The dependence on the solvent's refractive index introduced by Young and co-workers apparently produces a more accurate V value. For our high-S and high-V cases, classical theory is more accurate for estimating rate constants than is the use of vibronic coupling theory outside the parameter ranges for which it was designed (*27*). It remains to be seen how well a similar approach using κ_{el} in Eq. 2 will work when V is smaller and the ET reaction is not strongly adiabatic.

REFERENCES AND NOTES

1. R. A. Marcus and N. Sutin, *Biochim. Biophys. Acta* **811**, 265 (1985).
2. N. Sutin, *Prog. Inorg. Chem.* **30**, 441 (1983).
3. P. F. Barbara, T. J. Meyer, M. A. Ratner, *J. Phys. Chem.* **100**, 13148 (1996).
4. J. Jortner and J. Ulstrup, *J. Chem. Phys.* **63**, 4358 (1975); J. Jortner and M. Bixon, *ibid.* **88**, 167 (1988); J. Cortes, H. Heitele, J. Jortner, *J. Phys. Chem.* **98**, 2527 (1994).
5. G. L. Closs and J. R. Miller, *Science* **240**, 440 (1988).
6. M. R. Wasielewski, *Chem. Rev.* **92**, 435 (1992).
7. C. Creutz, *Prog. Inorg. Chem.* **30**, 1 (1983); K. Prassides, Ed., *Mixed Valency Systems: Applications in Chemistry, Physics, and Biology* (NATO ASI Ser., Kluwer Academic, Dordrecht, Netherlands, 1991).
8. N. S. Hush, *Prog. Inorg. Chem.* **8**, 391 (1967); *Coord. Chem. Rev.* **64**, 135 (1985).
9. There are significant problems in determining values of V for ET within photoexcited asymmetrical systems, which do have a transient to follow. Electronic mixing of the IV state with locally excited states is allowed, and both CT band intensities and back-ET rate constants are greatly affected by them. Good agreement with theory is obtained, but neither k_{et} nor ε_{max} relates to the V shown in Fig. 1 (*10*). For example, the V values estimated for Paddon-Row's σ-bridged aromatic, cyanoolefin systems (*11*), which appeared to be the best-documented series of such measurements available, have recently been retracted, with the statement that both the CT band intensities and back-ET rate constants are basically determined by mixing of the IV state with locally excited states (*10*).
10. M. Bixon, J. Jortner, J. W. Verhoeven, *J. Am. Chem. Soc.* **116**, 7349 (1994).
11. N. S. Hush *et al.*, *Chem. Phys. Lett.* **117**, 8 (1985); H. Oevering *et al.*, *J. Am. Chem. Soc.* **109**, 3258 (1987); M. N. Paddon-Row *et al.*, *J. Phys. Chem.* **92**, 6958 (1988); H. Oevering, J. W. Verhoeven, M. N. Paddon-Row, J. M. Warman, *Tetrahedron* **45**, 4751 (1989); J. M. Warman *et al.*, *J. Phys. Chem.* **95**, 1979 (1991); J. Kroon, J. W. Verhoeven, M. N. Paddon-Row, A. M. Oliver, *Angew. Chem. Int. Ed. Engl.* **30**, 1358 (1991); J. Kroon *et al.*, *J. Phys. Chem.* **97**, 5065 (1993).
12. S. F. Nelsen, H. Chang, J. J. Wolff, J. Adamus, *J. Am. Chem. Soc.* **115**, 12276 (1993).
13. S. F. Nelsen, J. Adamus, J. J. Wolff, *ibid.* **116**, 1589 (1994).
14. S. F. Nelsen, M. T. Ramm, J. J. Wolff, D. R. Powell, *ibid.* **119**, 6863 (1997).
15. S. F. Nelsen, R. F. Ismagilov, D. R. Powell, *ibid.* **118**, 6313 (1996).
16. ———, *ibid.*, in press.
17. C. Creutz, M. D. Newton, N. Sutin, *J. Photochem. Photobiol. A* **82**, 47 (1994).
18. S. F. Nelsen, *J. Am. Chem. Soc.* **118**, 2047 (1996).
19. I. R. Gould *et al.*, *Chem. Phys.* **176**, 439 (1993).
20. We obtained κ_{el} using equations 36 through 38 in (*2*), with the $h\nu_v = 800\ cm^{-1}$ used here.
21. Knowledge of ΔG° is far from sufficient to calculate k_{et} for a nonadiabatic case, where k_{et} is controlled by Franck-Condon factors, which in principle require a detailed knowledge of the contribution of all of the many modes (*22*) making up λ_v for their calculation. The adiabatic nature of our systems makes calculation of k_{et} far simpler than the case for nonadiabatic ET systems.
22. S. K. Doorn and J. T. Hupp, *J. Am. Chem. Soc.* **111**, 1142 (1989); A. B. Meyers, *Chem. Rev.* **96**, 911 (1996).
23. The quartic-augmented surfaces produce a value of λ that is slightly higher than E_{op} but closer to E_{op} than the λ obtained using vibronic coupling theory fits to the same data. $E_{1,min}$ is slightly more negative than $-V^2/\lambda$ obtained using parabolas, but the effect on $E_{1,max}$ is much larger. Use of $\Delta G^* \cong (\lambda/4)(1 + C/4)/(1 + C) - V + V^2/\lambda$ provides an excellent approximation to ΔG^* for our compounds.
24. We have evidence that the dielectric continuum approximation is a rather poor assumption for all of these compounds except for $\mathbf{1}^+$, and that it overestimates λ_s. Use of a smaller λ_s would make $(\lambda_v/\lambda)^{1/2}$ even closer to 1 and would not affect the estimation of $k_{et,ad}$ significantly.
25. The $n^{-1/2}$ dependence for V described in (*19*) arises from the use of a refractive index factor $f(n) = n^3$ in an expression for radiative emission (their equation 3). More recently, these authors have suggested using a more complex expression, $f(n) = n[(n^2 + 2)/3]^2$ [equation 3b of (*26*)], which they note produces a contribution that is about 11% smaller with a virtually identical dependence on n; they also indicate in appendix 3 (*26*) that the true dependence may be even more complicated.
26. I. R. Gould, R. H. Young, L. J. Mueller, A. C. Albricht, S. Farid, *J. Am. Chem. Soc.* **116**, 8188 (1994).
27. Because use of only λ and C fits the experimental IV-CT band shape, it is hardly surprising that, when the same data are fit using the three parameters of vibronic coupling theory (λ_s, λ_v, and $h\nu_v$), strong interaction of these parameters results (*14*, *16*). In studies using vibronic coupling theory to fit optical band shape or plots of k_{et} and ΔG°, it is usually stated that all three parameters were determined from the fit. This determination cannot be made for any of our data because of the strong interaction of these parameters.

28. We thank the National Science Foundation for partial financial support of this work through grant CHE-9417946 and R. Young, for extended discussions.

16 July 1997; accepted 22 September

Structure of the Carboxyl-Terminal Dimerization Domain of the HIV-1 Capsid Protein

Theresa R. Gamble,* Sanghee Yoo,* Felix F. Vajdos, Uta K. von Schwedler, David K. Worthylake, Hui Wang, John P. McCutcheon, Wesley I. Sundquist,† Christopher P. Hill†

The carboxyl-terminal domain, residues 146 to 231, of the human immunodeficiency virus–1 (HIV-1) capsid protein [CA(146–231)] is required for capsid dimerization and viral assembly. This domain contains a stretch of 20 residues, called the major homology region (MHR), which is conserved across retroviruses and is essential for viral assembly, maturation, and infectivity. The crystal structures of CA(146–231) and CA(151–231) reveal that the globular domain is composed of four helices and an extended amino-terminal strand. CA(146–231) dimerizes through parallel packing of helix 2 across a dyad. The MHR is distinct from the dimer interface and instead forms an intricate hydrogen-bonding network that interconnects strand 1 and helices 1 and 2. Alignment of the CA(146–231) dimer with the crystal structure of the capsid amino-terminal domain provides a model for the intact protein and extends models for assembly of the central conical core of HIV-1.

The 26-kD capsid protein (CA) performs essential roles both early and late in the life cycle of HIV. Capsid is initially translated as the central region of the 55-kD Gag polyprotein, where it functions in viral assembly (*1*, *2*) and in packaging the cellular protein cyclophilin A (*3*). As the virus buds, Gag is processed by the viral protease to produce three discrete new proteins—matrix, capsid, and nucleocapsid—as well as several smaller peptides. After capsid has been liberated by proteolytic processing, it rearranges into the conical core structure that surrounds the viral genome at the center of the mature virus (*1*, *4*). Genetic analyses have revealed that capsid also performs essential roles as the virus enters,

Department of Biochemistry, University of Utah, Salt Lake City, UT 84132, USA.

*These authors contributed equally to this report.
†To whom correspondence should be addressed.

uncoats, and replicates in the new host cell (*2*), although these early steps in viral replication are currently poorly understood.

Capsid is composed of two distinct domains (*2*, *5*, *6*). The NH_2-terminal domain (residues 1 to 145) contributes to viral core formation and binds cyclophilin A. In contrast, the COOH-terminal domain is required for capsid dimerization (*5*, *6*), Gag oligomerization (*7*), and virion formation (*2*). This domain also contains a conserved stretch of 20 amino acids, termed the MHR, that is found in all known onco- and lentiviruses and in the yeast retrotransposon Ty-3 (*8*, *9*) (Fig. 1). As would be expected from this conservation, the MHR is essential for viral replication, although its precise functions remain unclear because different MHR mutations block viral replication at distinct stages, including assembly (*9*–*12*), maturation (*10*, *12*), and early steps of infectivity (*9*, *11*, *12*).

In an effort to understand the structural basis for capsid dimerization and MHR conservation, we determined the crystal structure of selenomethionine-substituted CA(151–231) at 1.7 Å resolution (*13*, *14*) and native CA(146–231) at 3.0 Å resolution (*15*) (Tables 1 and 2). These two structures are similar, and therefore only CA(151–231) is described in detail, except where the capsid dimer interface is discussed (see below). The CA(151–231) model includes the initiator Met^{151} residue (which replaces Leu^{151} in the native sequence), capsid residues 152 to 220, and 63 ordered water molecules. The last 11 residues of capsid were not visible in electron density maps and have been omitted from the model. The R value is 22.9% and the R_{free} value is 27.4%.

Each CA(151–231) molecule is composed of an extended strand followed by four α helices (Fig. 2A) (*16*). The domain adopts an ovoid fold with overall dimensions of 27 Å by 29 Å by 38 Å. The MHR forms a compact strand-turn-helix motif that packs against the COOH-terminal end of helix 2 (Fig. 2B). The four most highly conserved MHR residues (Gln^{155}, Gly^{156}, Glu^{159}, and Arg^{167}) form a hydrogen-bonding network that stabilizes this structure and links it to helix 2. Within the network, the Gln^{155} side chain bridges the NH_2-terminus of helix 1 and the COOH-terminus of helix 2. The side chains of Glu^{159} and Arg^{167} form hydrogen bonds to main-chain amide and carbonyl groups across the MHR loop and also form a buttressing salt bridge. Finally, both main chain heteroatoms of Gly^{156} also form a hydrogen bond across the loop. The conservation of Gly^{156} can be explained because larger side chains would clash sterically with the Glu^{159} Cβ atom, particularly as the conserved Pro^{157} restricts the local chain flexibility. The structure also explains the observation that conservative mutations of invariant capsid MHR residues (for example, Gln^{155} → Asn or Glu^{159} → Asp) block HIV-1 replication (*10*), because even these single carbon mutations are expected to disrupt the hydrogen-bonding network that defines the MHR fold.

A series of conserved hydrophobic residues (Phe^{161}, Tyr^{164}, Val^{165}, Tyr^{169}, and Leu^{172}) are also essential for MHR function

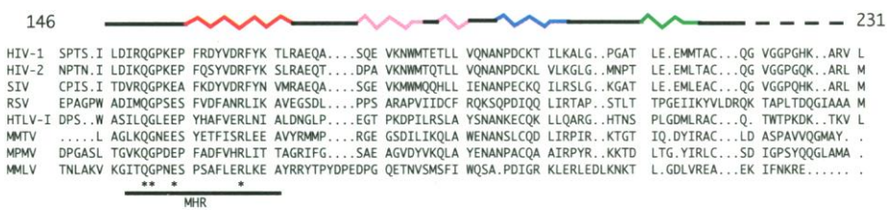

Fig. 1. HIV-1 CA(151–231) secondary structure and sequence alignment with representative onco- and lentiviruses. Disordered residues are indicated with a dashed line. The MHR is indicated with a bar, and invariant residues in this alignment are denoted with asterisks. Secondary structures are color-coded as follows: red, helix 1 (residues 11 to 24); magenta, helix 2 (29 to 42); cyan, helix 3 (46 to 54); green, helix 4 (61 to 67). SIV, simian immunodeficiency virus; RSV, Rous sarcoma virus; HTLV-I, human T cell leukemia virus–type I; MMTV, mouse mammary tumor virus; MPMV, Mason-Pfizer monkey virus; MMLV, Moloney murine leukemia virus. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

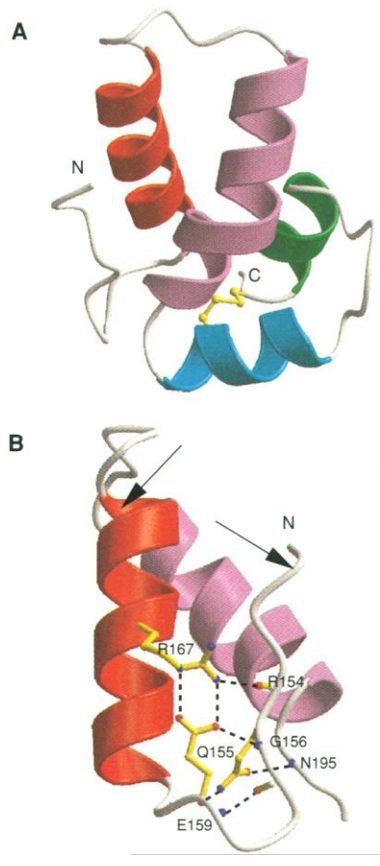

Fig. 2. (**A**) Structure of the CA(151–231) monomer. Color code is the same as in Fig. 1, with the disulfide bond between Cys^{198} and Cys^{218} shown in yellow. Residues 156 to 158 form a type I β-turn and residues 206 to 209 form a type II β-turn. (**B**) Expanded view of the MHR. The first and last residues of the MHR (residues 153 and 172) are indicated by arrows. The hydrogen-bonding interactions of the four most conserved MHR residues (Gln^{155}, Gly^{156}, Glu^{159}, and Arg^{167}) are shown explicitly. The following apparent hydrogen-bonding interactions (donor-acceptor distance of <3.1Å) are illustrated: Glu^{159} N···Gly^{156} O, Gln^{155} Nε2···Glu^{159} O, Asn^{195} N···Gln^{155} Oε1, Gly^{156} N···Glu^{159} Oε1, Arg^{167} Nη1···Glu^{159} Oε1, Arg^{167} Nη1···Arg^{154} O, and Arg^{167} Nε···Glu^{159} Oε2 (Figs. 2, 4, and 5 were made with MOLSCRIPT and RASTER 3D).

Table 1. Diffraction data. Parentheses denote highest shells of resolution.

Data set	Wavelength (Å)	Resolution (Å)	Unique reflections (N)	Redundancy*	Completeness (%)	R_{sym}† (%)
High res.‡	1.23980	25.0 to 1.7	9307	4.7	99.6 (98.6)	4.8 (28.0)
λ_1§	0.97907	15.0 to 2.1	8787	3.4	99.4 (96.0)	4.3 (14.7)
λ_2§	0.97873	15.0 to 2.1	8792	3.4	99.5 (97.3)	4.4 (14.9)
λ_3§	0.93000	15.0 to 2.1	8789	3.5	99.4 (99.8)	4.1 (9.9)
CA(146–231)‖	0.97800	20.0 to 3.0	2172	4.8	98.9 (100)	4.2 (21.5)

*Ratio of the total number of measurements to the number of unique reflections. †$R_{sym} = 100 \times \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{hi}$, where h are reflection indices, I_{hi} are intensities of symmetry-redundant reflections, and $\langle I_h\rangle$ is the mean intensity. ‡Selenomethionine-substituted CA(151–231) native data were collected at SSRL, beamline 1-5. §Selenomethionine-substituted CA(151–231) MAD data were collected at NSLS, beamline X12C. ‖CA(146–231) data were collected at NSLS, beamline X12B.

Table 2. Refinement statistics for CA(151–231). The model has 70 amino acid residues and 63 water molecules.

R value* (%)	22.9
R_{free} value† (%)	27.4
Resolution (Å)	6.0 to 1.7
rmsd, bond lengths (Å)	0.008
rmsd, bond angles (degrees)	1.415
$\langle B \rangle$ (Å^2) protein	19.4
$\langle B \rangle$ (Å^2) water	30.0
$\langle I/\sigma \rangle$	15.0

*R value = $100 \times \Sigma_h |F_{obs(h)} - F_{calc(h)}| / \Sigma_h F_{obs(h)}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes. †Ten percent of the reflections were reserved for the R_{free} value calculations.

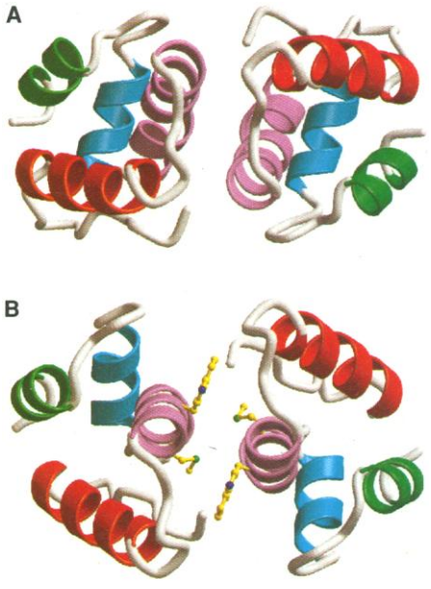

Fig. 4. (**A**) Structure of the CA(151–231) dimer. View is along the dyad axis, that is, from the top of Fig. 2A. (**B**) CA(146–231) dimer with the side chains of Trp[184] and Met[185] shown explicitly. Color code is the same as in Figs. 1 and 2.

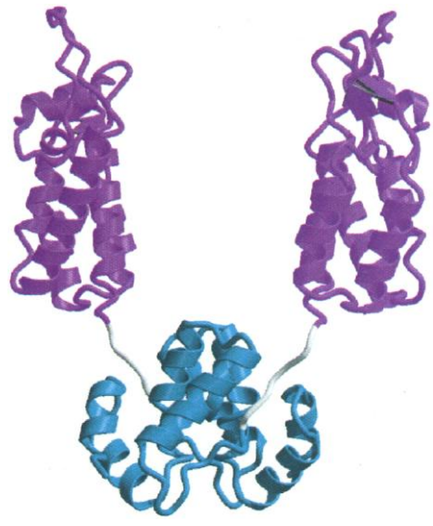

Fig. 5. Model for the intact HIV-1 capsid dimer. The CA(146–231) dimer (cyan) is shown covalently linked to the CA(1–151) domain (*23*). The five disordered residues that connect the two domains can be modeled to allow ~90° range of relative rotational orientation for the two domains about the vertical twofold axis.

(*8*, *9*). As predicted (*17*), these residues lie on one face of helix 1, where they contribute to the hydrophobic core of the protein. Thus, all of the conserved MHR residues perform critical structural roles. The remarkable conservation of the 20-residue MHR within the highly variable Gag polyprotein suggests that the MHR structure mediates an essential interaction with a highly conserved binding partner, such as a cellular factor or an invariant segment of a viral protein.

Dimerization of the intact capsid protein and the two COOH-terminal domain fragments was quantitated by equilibrium sedimentation. Native capsid exhibited a monomer-dimer equilibrium that was best fit with a dissociation constant K_d = 18 ± 1 μM (Fig. 3A), in good agreement with previous studies (*6*, *18*). The CA(146–231) protein dimerized with nearly the same affinity as the full-length protein (K_d = 10 ± 3 μM; Fig. 3B). Surprisingly, however, the slightly shorter CA(151–231) protein did not dimerize significantly, even at concentrations up to 100 μM (*19*). We therefore conclude that CA residues 146 to 151 are necessary for forming the high-affinity capsid dimer interface.

Both CA(146–231) and CA(151–231) formed dimers in the crystal through the packing of helix 2 across crystallographic twofold axes. The two dimers differ, however, by an apparent ~30° rotation of individual monomers (Fig. 4). The different dimerization geometries result from additional intermolecular contacts made by residues 146 to 151 (*15*). Our equilibrium sedimentation experiments indicate that the CA(146–231) dimer represents the structure of the high-affinity dimer interface of the intact protein (Fig. 3). This interface is created by the parallel packing of helix 2 against its symmetry-related mate, creating a hydrophobic core that includes residues Val[181], Trp[184], Met[185], and Leu[189]. The interface does not use MHR residues, which indicates that the MHR and dimerization motif are distinct entities that likely perform different functions.

We used site-directed mutagenesis to confirm that the dimer interface of the COOH-terminal CA(146–231) domain is also the principal dimer interface of the intact HIV-1 CA. Residue Met[185], which is in the center of the CA(146–231) dimer interface, was mutated to Ala in the context of the intact CA. This mutant protein

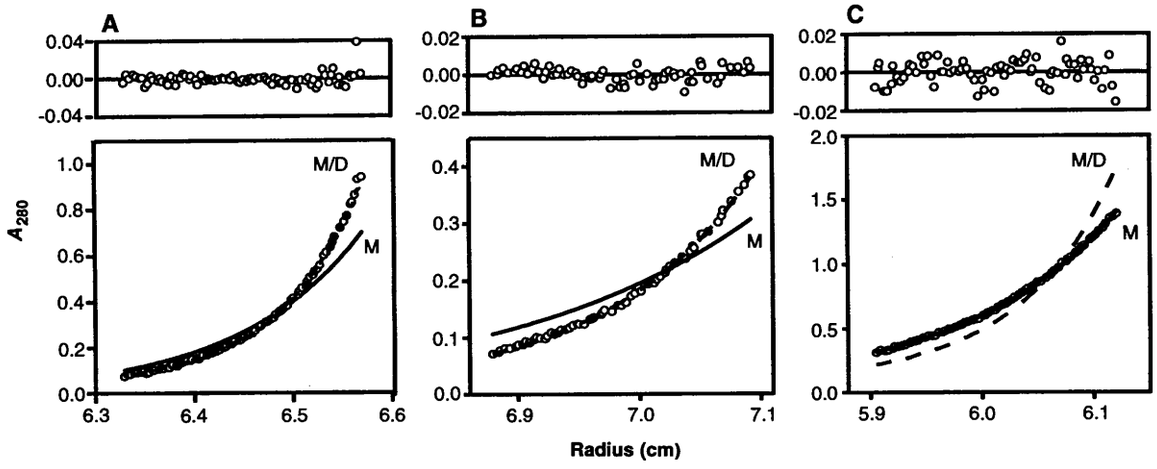

Fig. 3. Representative sedimentation profiles of (**A**) intact CA, (**B**) CA(146–231), and (**C**) CA(M185A) mutant protein. The analysis reveals that intact CA and CA(146–231) dimerize with similar affinities (K_d = 18 ± 1 and 10 ± 3 μM, respectively), but that the CA(M185A) mutation abolishes dimerization. Theoretical curves for monomer (M) and monomer/dimer (M/D) distributions are shown assuming K_d = 18 μM (A and C) or 10 μM (B). A_{280}, absorbance at 280 nm.

(M185A) was monomeric at all concentrations tested (Fig. 3C) (*20*). A second dimer interface mutation, $Trp^{184} \rightarrow Ala$ (W184A), similarly abolished capsid dimerization. Thus, helix 2 forms the principal dimer interface of the intact capsid protein.

We also tested the effect of the capsid M185A mutation on viral replication. This mutation reduced, but did not eliminate, the production of HIV-1(NL4–3) viral particles from 293T human embryonic kidney cells (*21*). The viral particles that were produced, however, were completely noninfectious. Thus, a point mutation that eliminates capsid dimerization in vitro also blocks viral replication in culture, which indicates that the crystallographically defined capsid dimer interface plays an essential role in HIV-1 replication.

The structures of both domains of HIV-1 CA are now known (*5*, *22*, *23*), and we have therefore examined possible models for the structure of the intact protein. The final ordered residue in both the crystal and nuclear magnetic resonance structures of capsid residues 1 to 151 [CA(1–151)] was Tyr^{145} (*5*, *23*). Thus, we modeled the intact CA by connecting the known CA(1–145) and CA(146–231) domains through a five-residue linker sequence corresponding to residues 146 to 150 (Fig. 5).

In the crystal structure of CA(1–151) in complex with cyclophilin A, the CA molecules associate into continuous planar strips that exhibit two distinct twofold symmetric interfaces (*23*). We speculate that the more extensive of these two NH_2-terminal domain interfaces (the β-hairpin interface) may cooperate with the COOH-terminal dimer interface described here to mediate higher order capsid protein assembly. Repetition of these two capsid dimer interfaces would create a strip of capsid molecules that could then wind up to create the core of HIV.

REFERENCES AND NOTES

1. H. G. Kräusslich, Ed., *Morphogenesis and Maturation of Retroviruses*, vol. 214 of *Current Topics in Immunology* (Springer-Verlag, Berlin, 1996).
2. J. B. M. Jowett, D. J. Hockley, M. V. Nermut, I. M. Jones, *J. Gen. Virol.* **73**, 3079 (1992); A. Von Poblotzki *et al.*, *Virology* **193**, 981 (1993); T. Dorfman, A. Bukovsky, Å. Öhagen, S. Höglund, H. G. Göttlinger, *J. Virol.* **68**, 8180 (1994); W.-H. Zhang, D. J. Hockley, M. V. Nermut, Y. Morikawa, I. Jones, *J. Gen. Virol.* **77**, 743 (1996); A. S. Reicin, Å. Öhagen, L. Yin, S. Höglund, S. P. Goff, *J. Virol.* **70**, 8645 (1996).
3. J. Luban *et al.*, *Cell* **73**, 1067 (1993); E. K. Franke, H. E. H. Yuan, J. Luban, *Nature* **372**, 359 (1994); M. Thali, *ibid.*, p. 363.
4. H. R. Gelderblom *et al.*, in *Membrane Interactions of HIV*, R. C. Aloia and C. C. Curtain, Eds. (Wiley-Liss, New York, 1992), pp. 33–55.
5. R. K. Gitti *et al.*, *Science* **273**, 231 (1996).
6. S. Yoo *et al.*, *J. Mol. Biol.* **269**, 780 (1997).
7. E. K. Franke, H. E. H. Yuan, K. L. Bossolt, S. P. Goff, J. Luban, *J. Virol.* **68**, 5300 (1994).
8. R. Patarca and W. A. Haseltine, *Nature* **318**, 390 (1985); J. W. Wills and R. C. Craven, *AIDS* **5**, 639 (1991).
9. K. J. Orlinsky, J. Gu, M. Hoyt, S. Sandmeyer, T. M. Menees, *J. Virol.* **70**, 3440 (1996).
10. F. Mammano, Å. Öhagen, S. Höglund, H. G. Göttlinger, *ibid.* **68**, 4927 (1994).
11. S. K. Strambio-de-Castillia and E. Hunter, *ibid.* **66**, 7021 (1992).
12. R. C. Craven, A. E. Leure-duPree, R. A. Weldon Jr., J. W. Wills, *ibid.* **69**, 4213 (1995).
13. DNA encoding an initiator methionine followed by residues 146 to 231 or 152 to 231 of HIV-1 capsid was amplified by the polymerase chain reaction, using the proviral pNL4-3 plasmid as a template. The amplified DNAs were subcloned into pET11a (Novagen) and sequenced, and the resulting constructs were expressed in *Escherichia coli* strains BL21(DE3) (native proteins) or B834(DE3) [selenomethionyl CA(151–231)] [F. W. Studier *et al.*, *Methods Enzymol.* **185**, 60 (1990)]. Proteins were purified by fractional ammonium sulfate precipitation followed by cation exchange and hydrophobic chromatographies.
14. Crystals of CA(151–231) were grown at 4°C in 4 μl of sitting drops containing a 1:1 mixture of protein solution [5.1 mM CA(151–231) in 10 mM tris (pH 8.0) and 2 mM 2-mercaptoethanol] and reservoir solution [20% polyethylene glycol (molecular weight 8000), 0.1 M sodium cacodylate (pH 7.4), and 0.2 M calcium acetate]. The space group was $P4_32_12$ ($a = b = 41.0$ Å, $c = 88.8$ Å) with one molecule per asymmetric unit. Before cryocooling, crystals were transferred into the reservoir solution containing 10% butanediol. All data were collected at 100 K and processed with DENZO and SCALEPACK (*24*). Data were collected to 2.1 Å resolution at three different wavelengths on a charge-coupled device (CCD) detector at beamline X12C of the National Synchrotron Light Source (NSLS). Structure determination was by the multiwavelength anomalous dispersion (MAD) method, with the λ_1 data treated as native and the λ_2, λ_3 data as derivatives [V. Ramakrishnan and V. Biou, *Methods Enzymol.* **276**, 538 (1997)]. All four selenium sites were located in difference Patterson and Fourier maps. Refinement of selenium atom parameters with MLPHARE [Collaborative Computing Project 4, *Acta Crystallogr.* **D50**, 760 (1994)] gave a mean figure of merit of 0.627. Phase refinement by solvent flattening and histogram shifting with DM (*ibid.*) gave a very clear map. A model consisting of the first 70 residues of CA(151–231) was built using O [T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard, *ibid.* **A47**, 110 (1991)] and was refined with X-PLOR (*25*) against 1.7 Å data that were collected on a MAR Imaging Plate detector at beamline 1-5 of the Stanford Synchrotron Radiation Laboratory (SSRL).
15. Crystals of CA(146–231) were grown at 4°C in 4 μl of sitting drops containing a 1:1 mixture of protein solution [2.1 mM CA(146–231) in 10 mM tris (pH 8.0) and 2 mM 2-mercaptoethanol] and reservoir solution (2 M ammonium sulfate). The space group was $I4_1$ ($a = b = 60.5$ Å, $c = 59.7$ Å) with one molecule per asymmetric unit. Crystals were transferred into a solution of 2 M ammonium sulfate and 20% glycerol, suspended in a loop, and cooled by plunging into liquid nitrogen. Data were collected at 100 K and processed with DENZO and SCALEPACK (*24*). Data were collected to 2.8 Å resolution on a CCD detector at beamline X12B of the NSLS. Structure determination was by molecular replacement with AMoRe [J. Navaza, *Acta Crystallogr.* **A50**, 157 (1994)] using the refined structure of CA(151–231) as the search model. The molecular replacement solution was clearly defined over a wide range of data resolution and integration radii. For example, with 8 to 3 Å data, the top rotation function solution had a correlation coefficient of 19.1% (next 13.0%); this rotation solution gave the top translation function solution with a correlation coefficient of 44.6% (next 36.4%). Rigid body refinement and overall *B* factor refinement gave an *R* value of 45.4% against all 8 to 3.0 Å data. The solution was further confirmed with simulated annealing omit maps (*25*). The CA(146–231) structure reported here is identical to that obtained following rigid body refinement, except for manual adjustment of the side chain torsion angles for Trp^{184}, which clearly differ in detail between the CA(151–231) and CA(146–231) structures. The CA(146–231) structure is currently being refined at 2.5 Å resolution.
16. A database search [L. Holmand and C. Sander, *J. Mol. Biol.* **233**, 123 (1993)] revealed that CA(151–231) most closely resembles the putative actin-binding subdomain of myosin [root-mean-square (rms) difference of 2.6 Å for 52 Cα atom pairs], with best overlap (rms difference, 1.0 Å) shown for helices 1 and 2 of CA(151–231) against myosin residues 524 to 536 and 545 to 551 [I. Rayment *et al.*, *Science* **261**, 58 (1993)]. Although the biological relevance of this similarity is not yet clear, the result is intriguing because Gag associates with the actin cytoskeleton [O. Rey, J. Canon, P. Krogstad, *Virology* **220**, 530 (1996)], and actin has been detected within HIV-1 at 200 molecules per virion [D. E. Ott *et al.*, *J. Virol.* **70**, 7734 (1996)].
17. C. B. Clish, D. H. Peyton, E. Barklis, *FEBS Lett.* **378**, 43 (1996).
18. S. Rosé *et al.*, *Proteins Struct. Funct. Genet.* **13**, 112 (1992).
19. T. R. Gamble, S. Yoo, F. F. Vajdos, U. K. von Schwedler, D. K. Worthylake, H. Wang, J. P. McCutcheon, W. I. Sundquist, C. P. Hill, data not shown.
20. The monomer-dimer equilibria of CA, CA(146–231), CA(151–231), and two mutant proteins [CA(W184A) and CA(M185A)] were analyzed by equilibrium sedimentation (*6*). The pure recombinant proteins were characterized by NH_2-terminal sequencing, mass spectrometry, and circular dichroism spectroscopy (to ensure that the mutations did not alter protein folding). Equilibrium centrifugation experiments were performed at rotor speeds of 20,000 rpm [CA and CA(W184A)] or 24,000 rpm [CA(146–231) and CA(151–231)] on a Beckman Optima XL-A. Equilibrium distributions of CA (from initial concentrations of 5.6, 9.8, and 15.8 μM) were globally fit to a monomer-dimer model, yielding a K_d of 18 ± 1 μM. Only the 9.8 μM data and associated residuals are shown in Fig. 3. Equilibrium distributions of CA(146–231) (from initial concentrations of 24.9, 65.4, and 81.7 μM) were globally fit to a monomer-dimer model, yielding a K_d of 10 ± 3 μM. Only the 24.9 μM data and associated residuals are shown in Fig. 3. Equilibrium distributions of CA(151–231) (from initial protein concentrations of 22.4 and 79.8 μM) were globally fit to simple monomer models, yielding an estimated protein mass of 9.6 ± 0.1 kD, in reasonable agreement with the mass of the CA(151–231) monomer (9.1 kD). Equilibrium distributions of CA(W184A) (initial protein concentrations of 7.5, 15.5, and 23 μM) and CA(M185A) (initial protein concentrations of 5.7, 11.8, and 16.9 μM) were globally fit to simple monomer models, yielding estimated protein masses of 25.5 ± 0.2 kD [CA(W184A)] and 25.5 ± 0.1 kD [CA(M185A)], in excellent agreement with their monomeric masses (25.5 kD). Only the 16.9 μM CA(M185A) data and associated residuals are shown in Fig. 3C.
21. DNA encoding the M185A mutation was subcloned into the full-length $HIV\text{-}1_{NL4\text{-}3}$ expression plasmid R9 [S. Swingler *et al.*, *J. Virol.* **71**, 4372 (1997)], which contains an SV-40 origin of replication. Viral particles were produced by transient calcium phosphate transfection of this plasmid into 293T human embryonic kidney cells. After 48 hours, particle production was reduced by 50 to 80% relative to a wild-type control R9 vector, as quantitated by both p24 enzyme-linked immunosorbent assay (DuPont) and a reverse transcriptase assay [S. P. Goff, P. Traktman, D. Baltimore, *J. Virol.* **38**, 238 (1981)]. Particle infectivity was assayed by normalizing mutant and wild-type virus for reverse transcriptase activity and by then performing viral growth curves in both CEM and SupT1 human T cell lines. No replication of the mutant virus could be detected in reverse transcriptase assays performed throughout a 2-week period; replication was therefore reduced by at least three orders of magnitude relative to wild-type virus. For a detailed mutational analysis of the capsid dimer interface, see U. K. von Schwedler *et al.*, in preparation.
22. C. Momany *et al.*, *Nature Struct. Biol.* **3**, 763 (1996).
23. T. R. Gamble *et al.*, *Cell* **87**, 1285 (1996).
24. Z. Otwinowski, in *Isomorphous Replacement and Anomalous Scattering*, W. Wolf, P. R. Evans, A. G. W. Leslie, Eds. (Daresbury Laboratory, Warrington, UK, 1991), pp. 80–86.
25. A. T. Brünger, *X-PLOR version 3.843, a System for X-ray Crystallography and NMR* (Yale University, New Haven, CT, 1996).
26. We thank M. Capel, C. Phillips, J. Phillips, R. Sweet,

and F. Whitby for assistance with data collection; V. Ramakrishnan and members of the Sundquist and Hill laboratories for critical comments on the manuscript; M. Martin for plasmid pNL4-3; D. Trono for plasmid R9; and J. Cassatt for support and encouragement. Supported by NIH grants RO1 AI40333 and RO1 AI43036 (W.I.S. and C.P.H.), the Lucille P. Markey Charitable Trust, and a postdoctoral fellowship from the Cancer Research Institute (T.R.G.). Coordinates (1am3) and diffraction data (r1am3sf) for CA(151–231) have been deposited in the Protein Data Bank (Brookhaven National Laboratory).

17 March 1997; accepted 23 September 1997

Metal Ion Chaperone Function of the Soluble Cu(I) Receptor Atx1

R. A. Pufahl, C. P. Singer, K. L. Peariso, S.-J. Lin, P. J. Schmidt, C. J. Fahrni, V. Cizewski Culotta, J. E. Penner-Hahn, T. V. O'Halloran*

Reactive and potentially toxic cofactors such as copper ions are imported into eukaryotic cells and incorporated into target proteins by unknown mechanisms. Atx1, a prototypical copper chaperone protein from yeast, has now been shown to act as a soluble cytoplasmic copper(I) receptor that can adopt either a two- or three-coordinate metal center in the active site. Atx1 also associated directly with the Atx1-like cytosolic domains of Ccc2, a vesicular protein defined in genetic studies as a member of the copper-trafficking pathway. The unusual structure and dynamics of Atx1 suggest a copper exchange function for this protein and related domains in the Menkes and Wilson disease proteins.

Although Cu is an essential cofactor for mitochondrial, cytosolic, and vesicular oxygen-processing enzymes (*1*), it can be toxic even at low concentrations. Cu(I) and Cu(II) ions can bind with high affinity to adventitious sites in partially folded proteins and catalyze auto-oxidation of lipids, proteins, and nucleic acids. To investigate the mechanisms by which cells overcome the dilemma of maintaining Cu availability while controlling deleterious reactivity of the free ions, we have determined the Cu chemistry and Cu-handling function of Atx1, an intracellular eukaryotic protein implicated in Cu trafficking. Our results indicate that Atx1 functions as a metal ion chaperone, a protein that protects and guides Cu(I) ions to activate target enzymes.

ATX1 is one of several genes in the Cu-dependent, high-affinity iron uptake pathway in yeast. These genes encode Ctr1, a Cu uptake protein in the plasma membrane; Ccc2, an intracellular membrane protein; and the multicopper oxidase Fet3 (*2–6*). Although the role of Fet3 in iron uptake is unclear, Cu loading into this enzyme occurs in a post-Golgi vesicle and is mediated by Ccc2 (*4*). The *ATX1* gene and its human homolog, *HAH1*, encode cytosolic proteins implicated in Cu trafficking to these Ccc2-containing vesicles (*7–9*). The Ccc2 protein in yeast and its human homologs, the Menkes disease protein (*10*) and Wilson disease protein (*11*), are members of the P-type adenosine triphosphatase (ATPase) cation transporter family (*12*) and are present in the membranes of secretory vesicles (*13*). Each transporter contains two or more Atx1-like cytoplasmic domains (Fig. 1), the functions of which are not known. The conserved MTCXXC sequence (X, any residue), a motif observed in several bacterial Hg(II) transport proteins (*14*), is thought to be a Cu binding site, although it does not correspond to known Cu(I) or Cu(II) sites in structurally characterized proteins. Establishing the Cu oxidation state and coordination environment in Atx1 should provide insight into its function and that of the homologous domains in its partner proteins.

Atx1 was expressed in *Escherichia coli* and purified as the apoprotein (*15*). Addition of $CuSO_4$ to apo-Atx1 in the presence of excess thiol reproducibly yielded a complex with a copper/protein ratio of 0.6 to 0.8 (*16*). Gel filtration experiments indicated that the predominant form of the protein was a monomer under these conditions, regardless of the presence or absence of Cu (*15*). The Cu oxidation state was investigated by electron paramagnetic resonance (EPR) spectroscopy and x-ray absorption spectroscopy (XAS). No EPR signal was observed at 77 K, suggesting the absence of a mononuclear Cu(II) site. XAS experiments (*17*) indicated that the bulk sample contained Cu(I). The Cu-Atx1 x-ray absorption near edge structure (XANES) spectrum exhibited a weak shoulder at 8984 eV, which is typical of Cu(I) and inconsistent with comparable edge features for Cu(II) compounds, the energies of which are 3 to 4 eV greater (Fig. 2A). The intensity of the 8984-eV transition varies with Cu(I) geometry, ranging from low for tetrahedral sites to high and well resolved for digonal sites (*18*). The observed transition is typical of those of trigonal Cu(I) model compounds (Fig. 2A). Mononuclear Cu(I) coordination complexes, however, are usually not stable in aqueous solution, and typically undergo auto-oxidation or disproportionation reactions to give Cu(II) (aqueous) and Cu(solid). In contrast, millimolar concentrations of Cu(I)-Atx1 are stable in air at neutral pH for at least 30 min (*16*), suggesting that the coordination environment in Atx1 stabilizes the Cu(I) state and suppresses disproportionation.

Proteins that stabilize Cu(I) form either polynuclear metal thiolate clusters, as in metallothionein and the transcription factors Ace1, CUP2, and AMT (*19*), or a constrained His_2Cys coordination environment, as in blue copper proteins (*20*). In contrast, extended x-ray absorption fine structure (EXAFS) measurements indicated that Atx1 stabilization of Cu(I) is achieved in a mononuclear site through an all-sulfur coordination environment (Fig. 2B). The high intensity and relatively high frequency of the EXAFS oscillations were typical of those observed for Cu-S interactions. The data could be modeled with a single three-sulfur (3S) shell with a Cu-S distance of 2.26 Å; however, the Debye-Waller factor, σ^2, was somewhat larger than expected (7×10^{-3} Å^2). This Cu-S distance is typical of three-coordinate Cu (*21–23*) and is 0.1 Å longer than that in two-coordinate Cu-thiolate complexes (*24*). Given that Atx1 contains only two conserved cysteine residues, the third ligand can be either a low-Z (atomic number) ligand (O or N), a methionine sulfur, or an exogenous thiol. A 2S + 1S fit with a shell of two sulfurs at 2.25 Å and a shell of one sulfur at 2.40 Å reproduced the data better than did the single-shell fit, and gave more reasonable σ^2 values (3×10^{-3} and 5×10^{-3} to 6×10^{-3} Å^2, respectively). Furthermore, attempts to model the data with a shell of 2S and a low-Z shell (1N) gave chemically unrealistic Debye-Waller factors for the N shell ($\sigma^2 \leq 0$).

The 2S + 1S result is unexpected for several reasons. The most common coordination environment for Cu(I) complexes is distorted four-coordinate, and there are few precedents among mononuclear Cu(I) complexes for the low-coordination number environ-

R. A. Pufahl, C. P. Singer, C. J. Fahrni, T. V. O'Halloran, Department of Chemistry and Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, IL 60208, USA.
K. L. Peariso and J. E. Penner-Hahn, Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA.
S.-J. Lin, P. J. Schmidt, V. Cizewski Culotta, Department of Environmental Health Sciences, Johns Hopkins University, Baltimore, MD 21205, USA.
*To whom correspondence should be addressed. E-mail: t-ohalloran@nwu.edu