JOURNAL OF VIROLOGY, June 2000, p. 5368–5372 Vol. 74, No. 11
0022-538X/00/$04.00+0

Structure-Function Studies of the Self-Assembly Domain of the Human Immunodeficiency Virus Type 1 Transmembrane Protein gp41

YONGKAI WENG,[1] ZHONGNING YANG,[2] AND CAROL D. WEISS[1]*

Center for Biologics Evaluation and Research, Food and Drug Administration,[1] and Laboratory of Structural Biology Research, National Institute for Arthritis and Musculoskeletal Diseases, National Institutes of Health,[2] Bethesda, Maryland 20892

Received 14 December 1999/Accepted 7 March 2000

The coiled-coil region of the human immunodeficiency virus type 1 transmembrane protein (gp41) makes up the interior core of the six-helix bundle structure of the gp41 self-assembly domain. We extended our previous study of this domain (Y. Weng and C. D. Weiss, J. Virol. 72:9676–9682, 1998) by analyzing 23 additional mutants at positions that lie at the interface of the interior core and outer helices. We found nine new functional mutants. For most mutants, the activity could be explained by the ability of the modeled mutants to stabilize the six-helix bundle structure. The present study provides insights into the envelope glycoprotein fusion mechanism and information for rational drug and vaccine design.

The envelope glycoprotein (Env) of the human immunodeficiency virus type 1 (HIV-1) mediates virus entry by fusing viral and cellular membranes. Membrane fusion is triggered when Env undergoes conformational changes while binding target cells. Env is synthesized as a fusion-incompetent precursor (gp160) that is processed into a receptor binding surface subunit (gp120) and a fusogenic transmembrane subunit (gp41). gp120 binding to CD4 and chemokine receptors activates gp41 fusion activity. Two heptad repeat regions in the gp41 ectodomain self-assemble into a thermostable six-helix bundle, consisting of a trimeric coiled-coil interior (N helix) with a three exterior helices (C helix) packed in the grooves of the trimer in an antiparallel manner (Fig. 1B and C) (1, 4, 13, 16, 17, 19). This six-helix bundle (hairpin structure) has been proposed to represent a fusion-active conformation of gp41.

In order to provide insights into the Env fusion mechanism and information for structure-based drug and vaccine design, we mutated residues at the interface of the N and C helices, in the e and g positions of the N-helix heptad repeat (Fig. 1A and 5A). Previously, we reported that nonconservative mutations in positions 570 and 577 from valine to glutamic acid (V570E) and glutamine to glutamic acid (Q577E), respectively, preserved fusion activity (18), but reanalysis of the V570E mutant (see below) shows that the mutation abolishes activity. In the three-dimensional structure, the 570 and 577 residues occupy positions in a hydrophobic cavity that has been proposed to be a good antiviral drug target (3). The present studies further assess the range of mutations tolerated in the 570 and 577 positions, as well as other equivalent positions along the interface between the N and C helices that lie outside the hydrophobic cavity.

Twenty-three Env mutants in six e or g positions in the N helix were generated by site-directed mutagenesis and verified by sequencing, as previously described (18). Two additional mutants in position 628 in the C helix were made to assess potential interactions with residue 577 in the hydrophobic cavity. All mutant Envs were analyzed for expression, processing, incorporation into virions (Fig. 2), and infectivity (Fig. 3), as previously described (18). One mutant, involving a leucine-to-lysine substitution in position 556 (L556K), lacked significant Env expression (data not shown) and consequently showed no Env incorporation into virions (Fig. 2A). Precursor processing of gp160 to gp120 and gp41 was greatly impaired for two mutants in the 628 position, involving a tryptophan-to-phenylalanine (W628F) or -alanine (W628A) substitution and slightly impaired for an arginine-to-glycine substitution in position 579 (R579G) (Fig. 2E). All other Env mutants showed Env expression, processing, and gp41 incorporation into virions at or near wild-type levels (Fig. 2B, C, and D).

The functions of the mutant Envs were assessed with a single-round infectivity assay by pseudotyping the mutant Envs with an Env-deficient viral genome containing a luciferase reporter gene (5), as previously described (18) (Fig. 3). Infectivity was determined by luciferase activity relative to that of the wild type. In the present study, we found nine new functional mutants (A558T, Q563A, Q563G, Q563R, V570I, Q577A, Q577D, Q577G, and Q577M), in addition to the Q577E mutant previously reported. The 563 and 577 positions, both e positions in the heptad repeat, tolerated multiple mutations. One functional mutant was identified in the 570 (V570I) and 558 (A558T) positions, which are e and g positions, respectively, in the heptad repeat.

The infectivity data provide insights into structural constraints on Env function. As expected, all mutants with impaired precursor cleavage (W628A, W628F, and R579G) lacked infectivity. These mutations likely affect folding of a prefusogenic Env conformation in a way that reduces accessibility of the gp120-41 junction to enzymatic cleavage. Except for the mutant that lacked Env expression (L556K), all other mutants are probably best interpreted in terms of their effects on a fusogenic conformation of Env. These mutants allowed gp41 incorporation into virions at or near wild-type levels and, therefore, probably do not disrupt the native conformation of Env. Presumably misfolded or nonnative Env would not be efficiently incorporated into virions.

For all mutants with normal gp41 incorporation into virions, we modeled each mutation in the six-helix bundle structure by

* Corresponding author. Mailing address: FDA/CBER, HFM-466, NIH Bldg. 29, Room 532, 29 Lincoln Dr., Bethesda, MD 20892-4555. Phone: (301) 402-3190. Fax: (301) 496-4684. E-mail: cdweiss@helix.nih.gov.

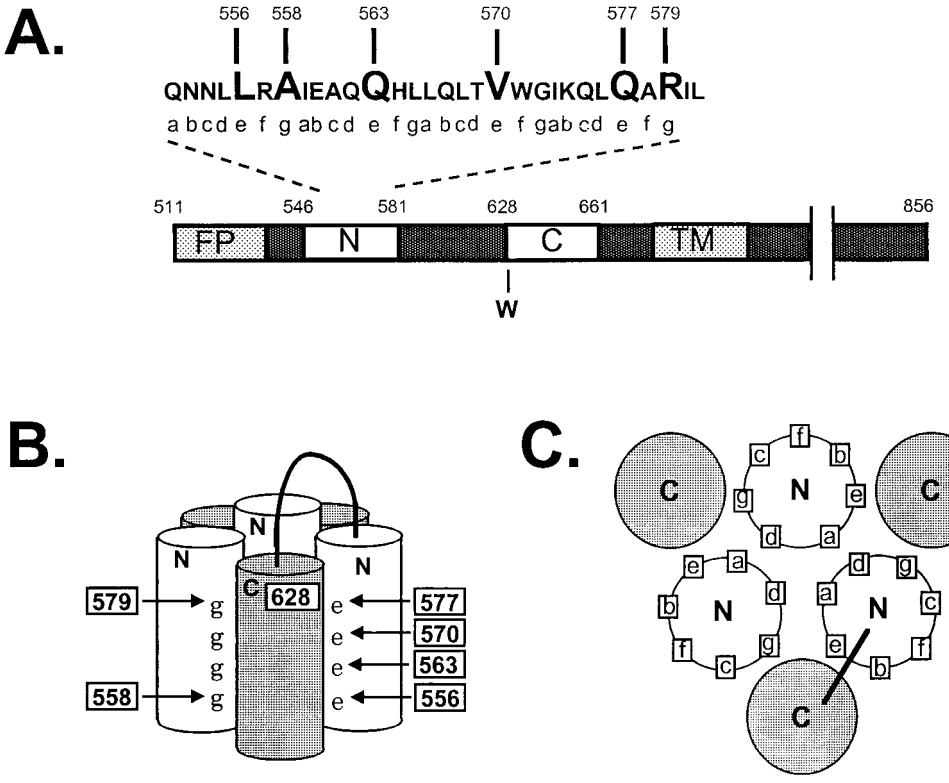

FIG. 1. Diagram of the gp41 self-assembly domain. (A) Linear representation of domains in gp41. FP, fusion peptide; N, heptad repeat motif located near the fusion peptide, containing sequences from DP-107 and N36; C, heptad repeat motif located near the transmembrane domain (TM) containing sequences from inhibitory C peptides (9, 11; C. Wild, T. Greenwell, and T. Matthews, Letter, AIDS Res. Hum. Retrovir. 9:1051–1053, 1993). Numbering corresponds to amino acid from the HXB2 clone of HIV-1 (Los Alamos National Laboratory, Los Alamos, N. Mex.). (B) Schematic representation of the self-assembly domain. N and C, heptad repeat regions that self-assemble to form the six-helix bundle structure (hairpin structure). e and g, positions of residues in the heptad repeat motif. Numbers indicate the positions of mutations. A loop connects the N and C helices from a gp41 monomer. The numbering is as described above and shows the antiparallel orientation of the N and C helices. (C) A cross-sectional view of the self-assembly domain, looking down from the connecting loop region, shows the orientation of heptad repeat residues.

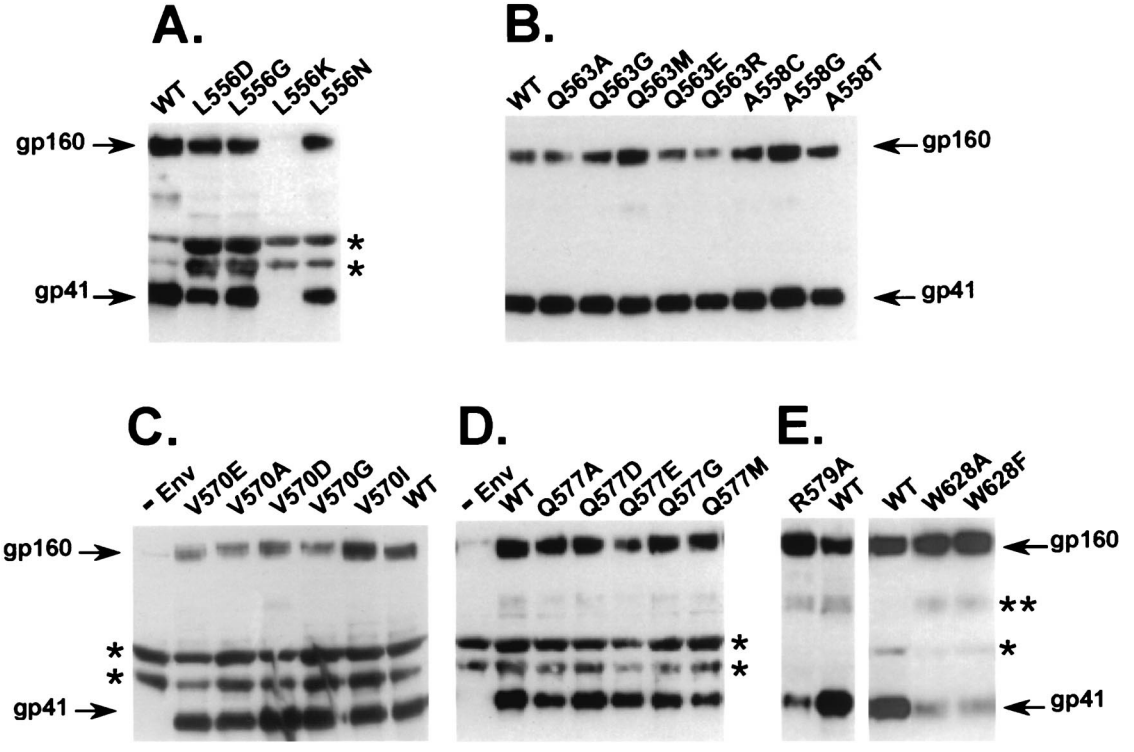

FIG. 2. Incorporation of gp41 into virions. (A to E) Virions were pelleted, lysed, and analyzed by immunoblotting with the Chessie 8 anti-gp41 monoclonal antibody, as previously described (18). Arrows indicate gp41 or gp160. −Env is a negative control. WT, wild-type sequence from the HXB2 clone of HIV-1. *, nonspecific bands; **, probable gp41 dimer. Gels are representative of at least three independent experiments.

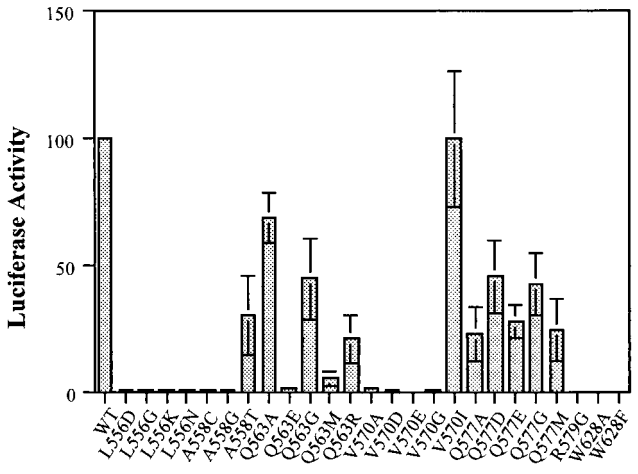

FIG. 3. Infectivity of mutant envelope glycoproteins. Mutant envelope glycoproteins were pseudotyped in an envelope-deficient HIV vector with a luciferase reporter gene as previously described (18). Relative infectivity was calculated by comparing the luciferase activity of mutant to wild-type virus × 100%. Results are an average of at least three independent experiments. Each error bar indicates 1 standard deviation.

using available crystallographic data (4, 13, 16, 17, 19). We analyzed interactions in the wild-type structure around each of the mutated sites to assess whether the activity of the mutants could be easily explained by the ability of the modeled mutation to fit in the six-helix bundle structure. For example, in the layers where residues 556, 558, and 570 are located, interhelical distances are smaller than the interhelical distances where residues 563 and 577 reside. The tighter interhelical packing probably imposes more constraints on residues 556, 558, and 570 and may account, in part, for the finding that fewer mutations are tolerated in these positions compared to in the 563 and 577 positions. Thus, in the tight hydrophobic pockets occupied by L556 and A558, as well as V570, substitutions with smaller residues such as glycine may leave the pockets unfilled and reduce interactions, which could destabilize the structure and account for the lack of infectivity of the L556G, A558G, V570G, and V570A mutants. In contrast, the large hydrophobic isoleucine substitution in position 570 (V570I) probably fills the pocket at least as well as the wild type and accounts for full infectivity (see Fig. 5C). Negatively charged residues may cause repulsion and dismantling of the hydrophobic pockets, accounting for the lack of infectivity of the L556D, V570D, and V570E mutants. Similarly, substitutions with polar residues, such as asparagine and cysteine, would not be favored in the hydrophobic surroundings and may explain the loss of activity for the L556N and A558C mutants. However, the substitution with threonine in position 558 (A558T) may retain fusion activity because of favorable hydrogen bonding with Q554.

Mutations in the 563 position behave differently from those in the 556, 558, and 570 positions, correlating with unique features in the layer where Q562 occupies a d position in the heptad repeat of the N helices. The large polar sidechain of glutamine, which is a relatively unusual residue in the a and d positions in heptad repeats, results in larger distances between both the N-N and N-C helices compared to other a or d layers in the six-helical bundle. This layer is stabilized by a hydrogen-bonding network between N and C helices, formed by the residue I559 carbonyl oxygen atom (backbone) and the Q562 nitrogen atom (side chain), as well as hydrogen-bonding networks within and between N helices, formed by Q562-Q563 side-chain interactions (Fig. 5B) and by residues Q562 with each other, respectively (see Fig. 4C in reference 19). Q563 side chains also interact with hydrophobic residues L565, I642, and I646 (Fig. 5B). Smaller substitutions with alanine and glycine in position 563 probably do not disturb Q562 hydrogen bonding, even though these residues do not fill the hydrophobic pocket as well as the wild type. Larger and more flexible substitutions at position 563, such as glutamic acid and methionine, could disrupt the Q562 side-chain-to-main-chain hydrogen bond and thus destabilize the helix bundle. The arginine substitution (Q563R) is probably too long to fit in the pocket, but may retain activity by rotating out of the pocket (Fig. 5B).

In the present studies, all mutants in the 577 (e) position in the hydrophobic cavity, including aspartic acid (Q577D), alanine (Q577A), glycine (Q577G), and methionine (Q577M), retained fusion activity, albeit at reduced levels (Fig. 3). The activity of multiple mutants in the 577 position suggests the presence of less-stringent packing constraints at this site, in agreement with increased interhelical distances observed in the structures (4, 13, 16, 17, 19). Previously, we speculated that the loss of hydrogen bonding between Q577 and W628 might account for the reduced fusion activity (18), based on structural data indicating that Q577E could form a salt bridge with an arginine in position 579 (R579) in an adjacent N helix and consequently lose hydrogen bonding with W628 (17) (Fig. 5E). Two high-resolution structures show potential for hydrogen bonding between Q577 and W628 (16, 17), but other structures suggest a stacking arrangement (4, 13, 19) (Fig. 5D). Our findings that diverse mutations without potential to form this hydrogen bond are tolerated in the 577 position argue against the hydrogen bonding between Q577 and W628 being required for Env function. However, the reduced fusion activity of the 577 mutants could be explained by either the loss of this hydrogen bond between N and C helices or less favorable stacking of the side chains of residues in the 577 and 628 positions.

To gain more information about the importance of hydrogen bonding between Q577 and Q628, we compared the Q577E mutant and the wild type for sensitivity to inhibition by a C-helix peptide containing W628, which is believed to inhibit Env by binding to the N helix. We speculated that if Q577E formed a salt bridge with R579 in an adjacent N helix and was

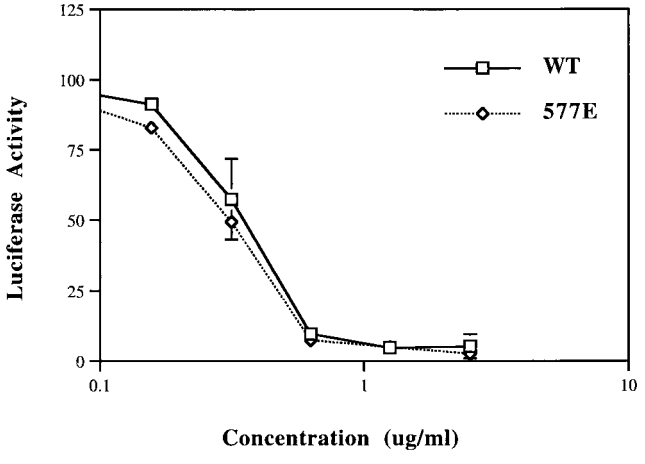

FIG. 4. Dose-response curve for inhibition of fusion by inhibitory C peptide containing W628. Synthetic C34 peptide (11) was added at the indicated final concentrations to cultures with pseudotyped virus stocks, as previously described (18). Percent inhibition was calculated by comparing luciferase activity at given peptide concentrations to control virus without peptide inhibitor × 100%. Results are an average of at least three independent experiments. WT, wild type. Each error bar indicates 1 standard deviation.

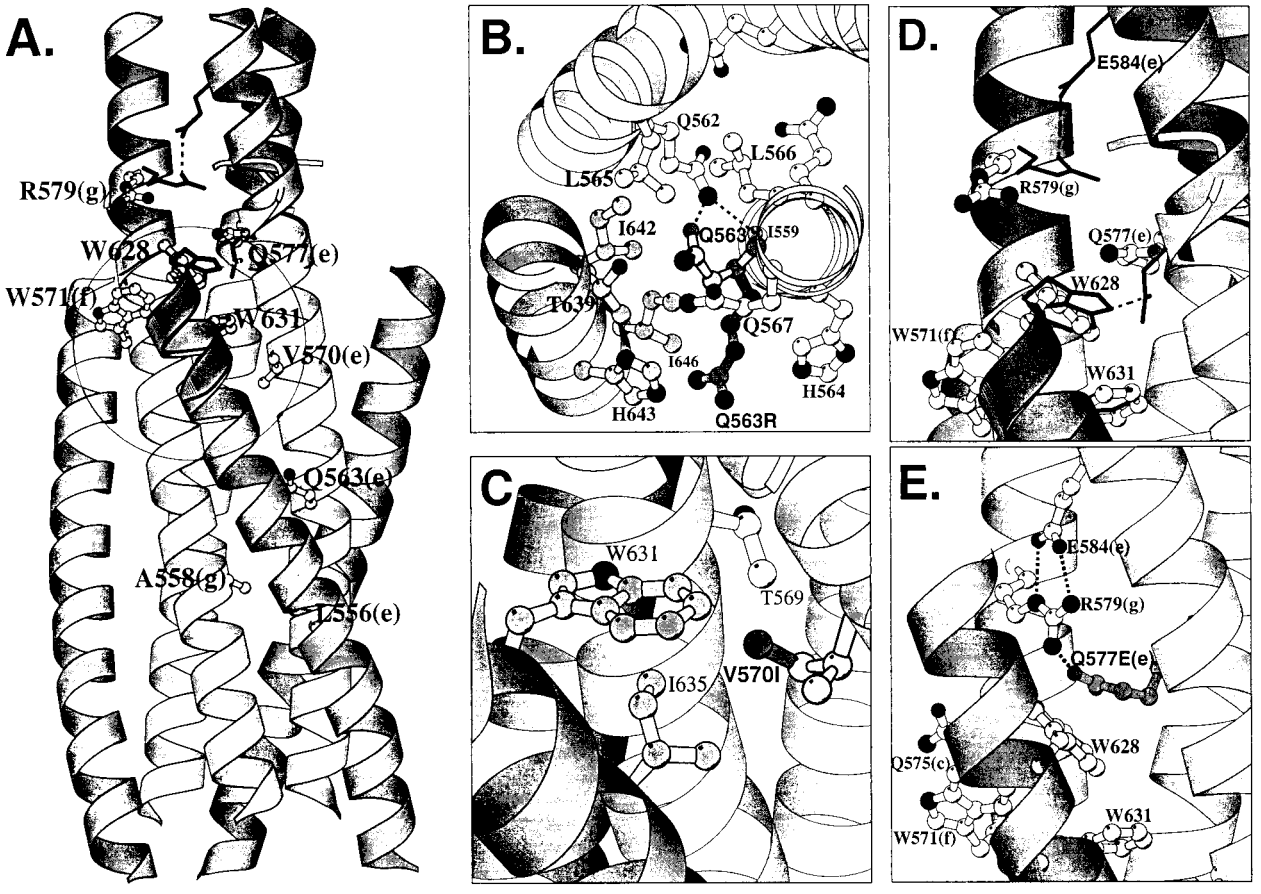

FIG. 5. Ribbon diagrams of the gp41 self-assembly domain generated with Molscript (10). Inner N helices and side chains are shown in light gray. Outer C helices are shown in dark gray. Dashed lines indicate potential hydrogen bonds. (A) Self-assembly domain shown in a longitudinal view, as in Fig. 1B. Structures from two coordinates are superimposed to highlight conformational differences. Lighter ribbon shading with ball-and-stick side chains uses the coordinates of Chan et al. (4), and darker ribbon shading with solid-line side chains near the N termini of helices use the coordinates of Weissenhorn et al. (17). A circle indicates the approximate position of the hydrophobic cavity (4). (B) Cross-sectional view of the self-assembly domain, as in Fig. 1C, highlighting molecular interactions affecting residue 563, using the coordinates of Chan et al. (4). Dotted lines highlight hydrogen bonds between residues Q562 and Q563 (side chain) and I559 (backbone oxygen atom) (4). The wild-type conformation of Q563 is shown in the center. A model of Q563R mutant is shown in the dark side chain. (C) Longitudinal view of the self-assembly domain highlighting molecular interactions affecting residue 570, using the coordinates of Chan et al. (4). The side chain of 570 is shown for the wild type (light gray) and the mutant modeled, V570I (dark gray). (D) Longitudinal view of the self-assembly domain, highlighting interactions between wild-type conformations of residues 577 in N helices and 628 in C helices. Ball-and-stick side chains use the coordinates of Chan et al. (4), and solid-line side chains use the coordinates of Weissenhorn et al. (17). (E) Longitudinal view of self-assembly domain modeling interactions between Q577E mutant in N helix (dark sidechain) and wild-type R579 in an adjacent N helix. Simian immunodeficiency virus gp41 coordinates of Yang et al. (19) are used because these coordinates are closely aligned to the coordinates of Chan et al. (4) and Weissenhorn et al. (17) and only require adjustment of the Q577E side chain.

not available for hydrogen bonding with W628 in a C helix, then the wild type might bind a C peptide inhibitor containing W628 better than the Q577E mutant and be more sensitive to inhibition. We found that Q577E did not significantly differ from the wild type in sensitivity to inhibition by the C-helix peptide (Fig. 4). Conversely, a study using mutant C peptide inhibitors showed that substitution of tryptophan with alanine in position 628 reduces the inhibitory activity of the peptide about fivefold, indicating that W628 in the C helix influences, but is not required for, peptide activity (3). These biological data, combined with structural data showing multiple orientations for Q577 and W628 (1, 4, 13, 16, 17, 19), suggest that there is some flexibility in the interactions of the C terminus of the C peptide with the N helices (Fig. 5D).

Attempts to generate mutants in the 579 position (g) were unsuccessful except for an arginine-to-glycine substitution (R579G). The R579G mutant showed impaired cleavage of the gp160 precursor, but to a much lesser extent than the 628 mutants. The high-resolution structures indicate that the arginine in position 579 (R579) in the wild-type structure has the potential to form a salt bridge with glutamic acid in position 584 in an adjacent N helix (Fig. 5E). R579 may also hydrogen bond to glutamine in position 575 of the same N helix as well as W628 in the C helix (Fig. 5E). The complete absence of fusion activity of the R579G mutant may not be completely accounted for by the impairment of gp160 processing and may reflect, in part, the loss of these inter- and intrahelical interactions from the glycine substitution. Also, we previously speculated that interactions between 579R in the N helix and 577E in an adjacent N helix might account for a slightly increased sensitivity of the 577E mutant to inhibition by an N-helix peptide (DP-107). However, other functional mutants in the 577 position, which cannot form a salt bridge with R579, showed no significant differences in their sensitivity to inhibition by the DP-107 peptide compared to Q577E (data not shown) and do not support this interpretation.

Other mutagenesis studies have shown that a variety of mutations in the gp41 ectodomain negatively affect gp41 fusion

activity (2, 6–8, 12, 15, 18). Analysis of an escape-mutant virus that is resistant to an inhibitory C peptide (DP-178) showed mutations in the N helix distinct from the mutations reported here (14). Our studies analyzed novel mutants to probe interactions at the interface of N and C helices. We found that diverse mutations, including nonconservative changes in hydrophobic cavities, are tolerated in several positions in the N helix. This surprising finding suggests the potential for the emergence of viruses that are resistant to agents targeting the highly conserved self-assembly domain. Nonetheless, most mutants were impaired in fusion activity, and the only mutant that had full wild-type activity, V570I, is represented in the unusual HIV-1 variant CPZGAB. The activity of most mutants could be readily explained by the ability of the modeled mutations to support interactions in the six-helix bundle structure, suggesting that the six-helix bundle structure is critical for Env fusion activity. These data provide insights into the Env fusion mechanism and information for rational drug and vaccine design.

We thank Ira J. Berkower and Eve de Rosny (Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Md.) and Paul Wingfield and Tim Mueser (National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Md.) for helpful discussions and critical reading of the manuscript. We also thank C. Craig Hyde (National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health) for support, Carl T. Wild (Biotech Research Labs, Gaithersburg, Md.) for providing N and C helix peptides, Anthony Sun (Rutgers University, Piscataway, N.J.) for technical assistance, and Michael Klutch (Center for Biologics Evaluation and Research) for DNA sequencing support.

REFERENCES

1. **Caffrey, M., M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, D. G. Covell, A. M. Gronenborn, and G. M. Clore.** 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. EMBO J. **17:**4572–4584.
2. **Cao, J., L. Bergeron, E. Helseth, M. Thali, H. Repke, and J. Sodroski.** 1993. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. J. Virol. **67:** 2747–2755.
3. **Chan, D. C., C. T. Chutkowski, and P. S. Kim.** 1998. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. Proc. Natl. Acad. Sci. USA **95:**15613–15617.
4. **Chan, D. C., D. Fass, J. M. Berger, and P. S. Kim.** 1997. Core structure of gp41 from the HIV envelope glycoprotein. Cell **89:**263–273.
5. **Chen, B. K., K. Saksela, R. Andino, and D. Baltimore.** 1994. Distinct modes of human immunodeficiency virus type 1 proviral latency revealed by superinfection of nonproductively infected cell lines with recombinant luciferase-encoding viruses. J. Virol. **68:**654–660.
6. **Chen, S. S.-L.** 1994. Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane protein gp41. J. Virol. **68:** 2002–2010.
7. **Chen, S. S.-L., C.-N. Lee, W.-R. Lee, K. McIntosh, and T.-H. Lee.** 1993. Mutational analysis of the leucine zipper-like motif of the human immunodeficiency virus type 1 envelope transmembrane glycoprotein. J. Virol. **67:** 3615–3619.
8. **Dubay, J. W., S. J. Roberts, B. Brody, and E. Hunter.** 1992. Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. J. Virol. **66:**4748–4756.
9. **Jiang, S., K. Lin, N. Strick, and A. R. Neurath.** 1993. HIV-1 inhibition by a peptide. Nature **365:**113.
10. **Kraulis, P. J.** 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. **24:**946–950.
11. **Lu, M., S. C. Blacklow, and P. S. Kim.** 1995. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. Nat. Struct. Biol. **2:**1075–1082.
12. **Lu, M., H. Ji, and S. Shen.** 1999. Subdomain folding and biological activity of the core structure from human immunodeficiency virus type 1 gp41: implications for viral membrane fusion. J. Virol. **73:**4433–4438.
13. **Malashkevich, V. N., D. C. Chan, C. T. Chutkowski, and P. S. Kim.** 1998. Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. Proc. Natl. Acad. Sci. USA **95:**9134–9139.
14. **Rimsky, L. T., D. C. Shugars, and T. J. Matthews.** 1998. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. J. Virol. **72:**986–993.
15. **Salzwedel, K., J. T. West, and E. Hunter.** 1999. A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. J. Virol. **73:**2469–2480.
16. **Tan, K., J. Liu, J. Wang, S. Shen, and M. Lu.** 1997. Atomic structure of a thermostable subdomain of HIV-1 gp41. Proc. Natl. Acad. Sci. USA **94:** 12303–12308.
17. **Weissenhorn, W., A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley.** 1997. Atomic structure of the ectodomain from HIV-1 gp41. Nature **387:** 426–430.
18. **Weng, Y., and C. D. Weiss.** 1998. Mutational analysis of residues in the coiled-coil domain of the human immunodeficiency virus type-1 transmembrane protein gp41. J. Virol. **72:**9676–9682.
19. **Yang, Z. N., T. C. Mueser, J. Kaufman, S. J. Stahl, P. T. Wingfield, and C. C. Hyde.** 1999. The crystal structure of the SIV gp41 ectodomain at 1.47 angstrom resolution. J. Struct. Biol. **126:**131–144.

ERRATUM

Structure-Function Studies of the Self-Assembly Domain of the Human Immunodeficiency Virus Type 1 Transmembrane Protein gp41

YONGKAI WENG, ZHONGNING YANG, AND CAROL D. WEISS

Center for Biologics Evaluation and Research, Food and Drug Administration, and Laboratory of Structural Biology Research, National Institute for Arthritis and Musculoskeletal Diseases, National Institutes of Health, Bethesda, Maryland 20892

Volume 74, no. 11, p. 5368–5372, 2000. Page 5369, Fig. 2E, first lane: "R579A" should read "R579G."