JOURNAL OF VIROLOGY, Mar. 1996, p. 2038–2043 Vol. 70, No. 3
0022-538X/96/$04.00+0

In Vivo Sequence Diversity of the Protease of Human Immunodeficiency Virus Type 1: Presence of Protease Inhibitor-Resistant Variants in Untreated Subjects

WALTER J. LECH,[1] GREG WANG,[1] YAN LI YANG,[1] YING CHEE,[1] KARIN DORMAN,[2] DEREK MCCRAE,[1] LAURA C. LAZZERONI,[3] JOHN W. ERICKSON,[4] JANET S. SINSHEIMER,[2] AND ANDREW H. KAPLAN[1,5]*

Departments of Medicine,[1] Biomathematics,[2] and Microbiology and Immunology,[5] University of California at Los Angeles, Los Angeles, California 90024; Department of Statistics, Stanford University, Stanford, California 94305[3]; and the Structural Biochemistry Program, PRI/Dyn Corporation, National Cancer Institute Frederick Cancer Research and Development Center, Frederick, Maryland 21702[4]

We have evaluated the sequence diversity of the protease of human immunodeficiency virus type 1 in vivo. Our analysis of 246 protease coding domain sequences obtained from 12 subjects indicates that amino acid substitutions predicted to give rise to protease inhibitor resistance may be present in patients who have not received protease inhibitors. In addition, we demonstrated that amino acid residues directly involved in enzyme-substrate interactions may be varied in infected individuals. Several of these substitutions occurred in combination either more or less frequently than would be expected if their appearance was independent, suggesting that one substitution may compensate for the effects of another. Taken together, our analysis indicates that the human immunodeficiency virus type 1 protease has flexibility sufficient to vary critical subsites in vivo, thereby retaining enzyme function and viral pathogenicity.

The protease of human immunodeficiency virus type 1 (HIV-1) plays a critical role in the viral life cycle. Mature virus particles are surrounded by a lipid envelope derived from the host cell and contain a centrally placed electron-dense core structure. The proteins of the core are translated as part of two precursors, Gag and Gag/Pol. During virus assembly, the Gag and Gag/Pol polyprotein precursors are cleaved by the viral protease, thereby liberating the mature structural and enzymatic proteins of the core (6). Inactivation of the protease results in the production of noninfectious, aberrantly assembled viral particles composed of unprocessed precursors (14, 15, 35).

Recently, a number of inhibitors of the HIV-1 protease have become available and have shown some initial promise in clinical trials (5, 30, 41, 46). Treatment with these compounds has been associated with marked declines in the number of copies of viral RNA per milliliter of plasma and increases in the number of $CD4^+$ cells in infected patients. Unfortunately, protease inhibitor-resistant variants of HIV-1 have appeared both during selection in cell culture and *in vivo* during treatment (5, 8, 11–13, 21, 24, 31, 33, 38, 42). Substitutions in the protease coding domain that are associated with reduced sensitivity to many of these agents have been identified. Further, in several instances, the structural basis for the reduced sensitivity has been determined by modeling these mutations into the available crystal structures of protease-inhibitor complexes (11, 13, 24). In addition, other mutations which appear to compensate for the deleterious effects of the resistance mutations have also been characterized in structural models as well as in the context of viral replication in cell culture (11).

Although the presence of resistant viruses may make therapy with this class of agents less effective, it is possible that the viruses carrying these mutations are attenuated and less virulent. Alternatively, resistant variants may acquire compensatory mutations elsewhere in the protein and thereby at least partially restore enzyme function. As part of an attempt to characterize the forces that drive the molecular evolution of the viral protease and define viral fitness, we have obtained 246 protease coding domain sequences from 12 HIV-infected persons, none of whom have received protease inhibitors. We found that a substantial number of these patients carry viral genomes with substitutions in the protease that are predicted to give rise to inhibitor-resistant enzymes. In addition, our data suggest that compensatory mutations predicted to enhance enzyme function may occur in vivo.

The 12 subjects were patients at the University of California at Los Angeles Center for AIDS Research and Education. Each subject was receiving zidovudine and had $CD4^+$ cell counts below 500 cells per mm^3. None of them had received a protease inhibitor. Between 15 and 24 clones were obtained from each subject (median, 21 per subject). Two subjects were excluded. We were unable to amplify the protease coding domain from one subject despite multiple attempts. Sequences from an additional subject yielded stop codons (at positions 6 and 42) for all 23 clones sequenced; this subject was excluded from the analysis. Blood was obtained from the study subjects, and between 1×10^6 and 10×10^6 lymphocytes were collected by centrifugation. Peripheral blood lymphocytes were disrupted with sodium dodecyl sulfate-proteinase K. The DNA was isolated from the cells with phenol-chloroform and extracted with ethanol according to published procedures (37). The provirus copy number was determined by using a quantitative DNA protocol described elsewhere (9).

The protease coding domain was amplified by using the following primers in a nested PCR protocol (incorporated restriction sites are indicated in lowercase letters): PR2.1849, 5′-GATGACAGCATGTCAGGGAGTA-3′, and PR2.2796, 5′-CTTCCCAGAAGTCTTGAGTTCT-3′ (outer primers), and PR2.1937, 5′-ATGctgcagAGAGGCAATTT-3′ (*Pst*I), and PR2.2716, 5′-GGCAAATActcgagTATTGT-3′ (*Xho*I) (inner

* Corresponding author. Mailing address: 37-121 CHS, UCLA School of Medicine, 10833 LeConte, Los Angeles, CA 90024. Phone: (310) 825-7225. Fax: (310) 825-3632. Electronic mail address: akaplan@medicine.medsch.ucla.edu.

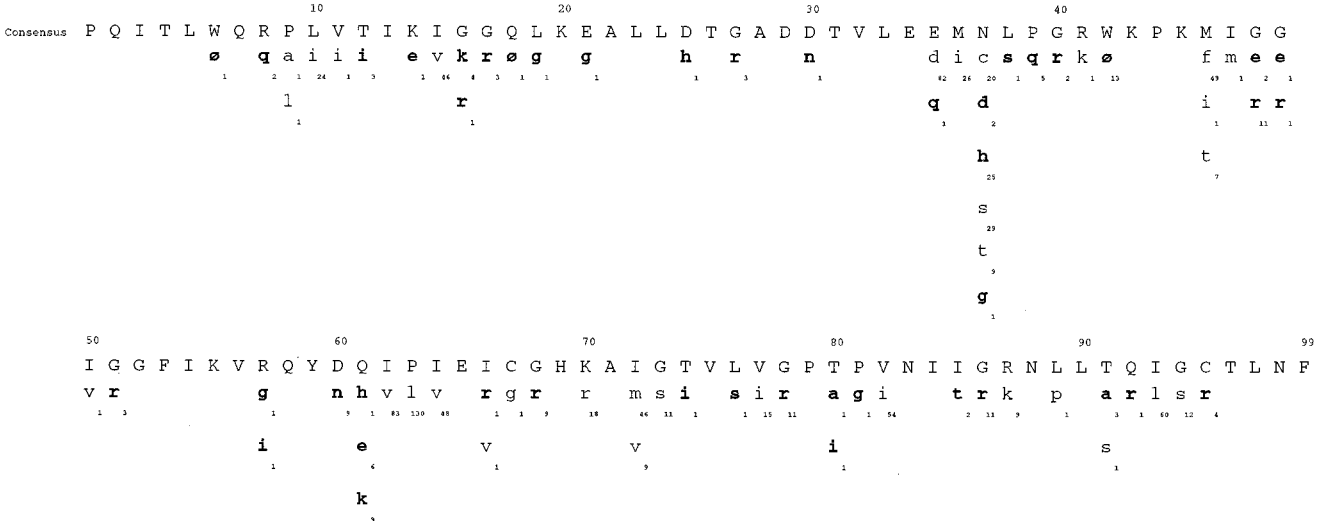

FIG. 1. Amino acid variation in 246 sequences of the HIV-1 protease from 12 patients. Amino acids were deduced from nucleotide sequences obtained by PCR amplification, cloning, and dideoxynucleotide sequencing of the protease open reading frame. The single-letter amino acid code is used, and ø's represent stop codons. The consensus sequence is listed on the top line. Variations from the consensus at each position are listed below the consensus amino acid; the number of times that each variant appears is denoted by the subscript number. Nonconservative substitutions are shown in boldface.

primers). The reactions were carried out in a solution containing final concentrations of 10 mM Tris-HCl (pH 8.3), 2 mM $MgCl_2$, 50 mM KCl, 0.2 mM (each) the deoxynucleoside triphosphates, and 2.5 U *Taq* polymerase in a final volume of 100 μl. Amplifications were performed in a Perkin-Elmer Thermal Cycler 9600. Synthesis of the first-round product consisted of 2 min at 96°C followed by 35 cycles of 20 s at 95°C, 20 s at 53°C, and 30 s at 72°C, with a final extension at 72°C for 6 min. One microliter of the first-round product was used as the template for the second-round amplification with PR2.1937 and PR2.2716 consisting of 2 min at 96°C followed by 35 cycles of 20 s at 95°C, 20 s at 52°C, and 30 s at 72°C, with a final extension at 72°C for 6 min. Purified DNA from the PCR products was inserted into the Bluescript vector (Gibco/BRL) by using the *Xho*I and *Pst*I restriction enzyme sites incorporated into the inner PCR primers. The entire protease coding domain was sequenced with Sequenase version 2.0 (United States Biochemical) on both DNA strands according to the manufacturer's recommendations. We have demonstrated previously that this protocol introduces approximately one substitution for every 1,800 bases of the protease coding domain amplified and sequenced (47).

The nucleotide sequences were translated by using the MAP subroutine of the University of Wisconsin Genetics Computer Group software package (OPENVMS, version 8.0; copyright 1994). The reading frame for each sequence was determined by comparison with a reference HIV-1 sequence (GenBank accession number K03455). Amino acid sequences were aligned by using the multiple sequence alignment subroutine PILEUP. Analyses were performed to evaluate whether substitutions at two positions are independent. This would imply that knowledge of the amino acid at one site does not impart additional information for prediction of the amino acid present at the second site. Let p_A be the marginal probability of observing amino acid A at position 1, and let q_B be the marginal probability of observing amino acid B at position 2. If the joint probability of observing amino acid A at position 1 and amino acid B at position 2 is $r_{AB} = p_A \times q_B$ for all possible pairs A,B of amino acids, then substitutions at positions 1 and 2 are independent. For three sites, mutual independence of the sites implies $r_{ABC} = p_A \times q_B \times s_C$. This definition can be extended to any number of amino acid positions. Hypothesis testing for independence used exact conditional methods for analyzing contingency tables. Monte Carlo simulation was used to determine the distribution of the chi-square statistic given the observed amino acid totals under the null hypothesis of independence (20).

Heterogeneity of the HIV-1 protease coding domain in vivo includes regions known to be critical for enzyme-substrate interactions. We examined the extent of protease sequence heterogeneity in HIV-infected individuals who had not received protease inhibitors. A total of 246 sequences were recovered from the viral DNA present in the peripheral blood lymphocytes of 12 subjects. In order to produce a representative collection of protease sequences, for each subject, with the exception of subject K, PCR was performed on a minimum of 120 copies of HIV DNA. For subject K, a sample of peripheral blood lymphocyte DNA containing 40 copies of HIV DNA was used.

We constructed a consensus sequence by using the entire database; the individual variations from the consensus are summarized in Fig. 1. There are a total of 947 substitutions. Overall, the heterogeneity at the nucleotide level as determined by pairwise comparison varied from 0.9997 to 0.9465 among the 12 subjects (data not shown). For the entire data set, the ratio of synonymous substitutions to nonsynonymous substitutions was 3.41. Most of the substitutions that we identified (84%) were conservative (Fig. 2). Further, the substitutions were clustered, sparing several regions. The amino acids that define the cleavage sites at either end of the protein are almost absolutely conserved. Although only the protease coding domain is shown, the amino acids immediately upstream and downstream of the cleavage sites that release the mature protease were also sequenced and are similarly conserved (data not shown). In addition, the three mutationally sensitive domains identified by Loeb et al. are also conserved (22). On the basis of crystallographic studies these domains correspond to the active-site loop (residues 22 to 33), which includes the Asp-Thr-Gly active-site triad characteristic of aspartic proteases, the flap region (residues 47 to 52), and the hydrophobic core of the molecule (residues 74 to 87). Sixteen of the clones encode enzymes that are obviously defective, including 15

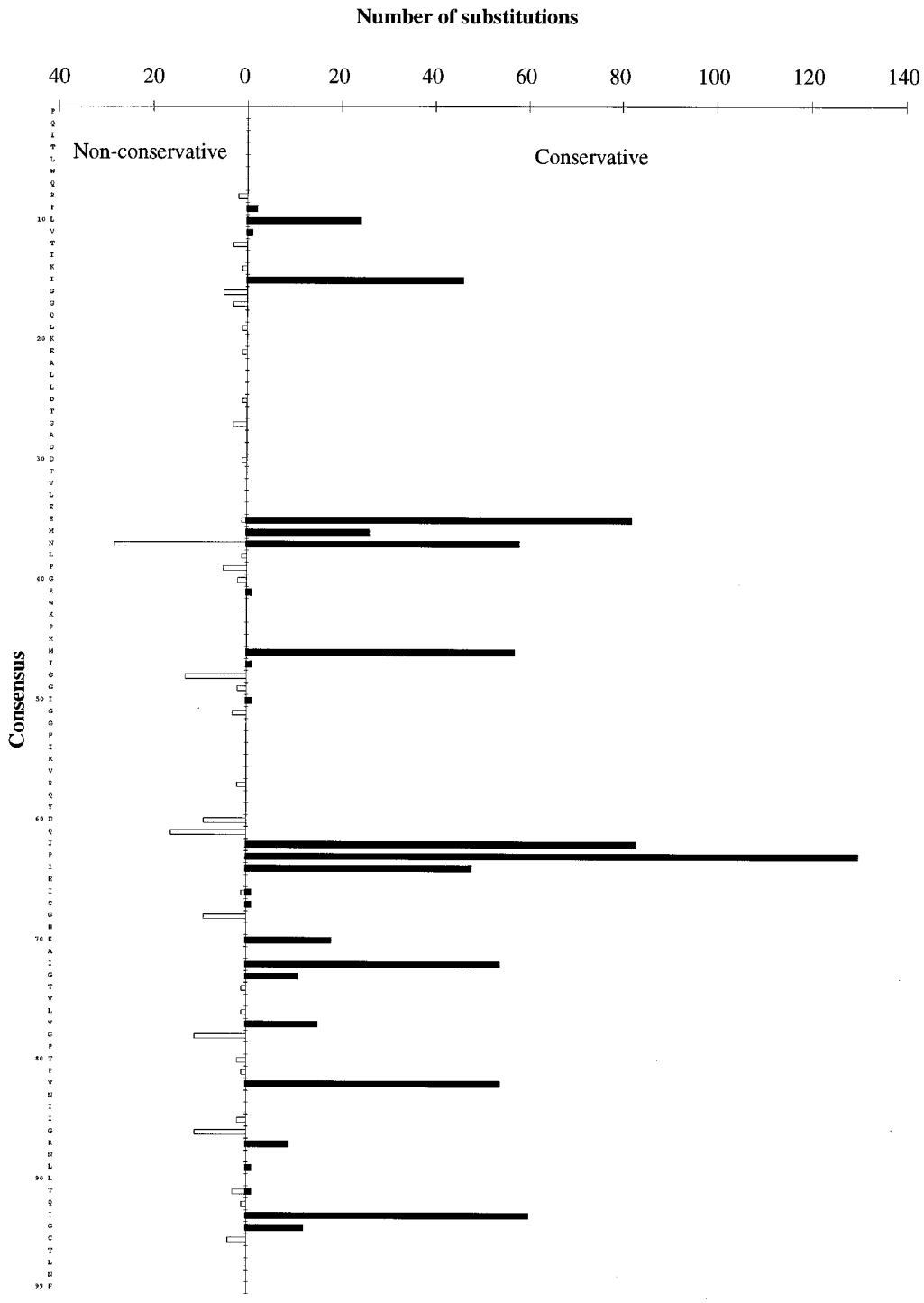

FIG. 2. Summary of conservative versus nonconservative substitutions in the HIV-1 protease. The total numbers of conservative and nonconservative substitutions relative to the consensus sequence are displayed. The consensus sequence is shown on the left. The total number of conservative substitutions at each position is shown by the length of the filled bar to the right of the line at 0. Nonconservative substitutions are shown by the open bars. The scale at the top shows the total number of substitutions at each position.

clones with stop codons and 1 with a mutation in the active-site aspartic acid (i.e., a mutation of D to H at position 25 [D25→H]).

A substantial number of the clones displayed variation in the amino acids that comprise the subsites of the protease implicated in substrate binding (Fig. 1) (for a review, see reference 46). There are 14 amino acids on both protease monomers whose side chains make up the two subsites (S1 and S1′) that interact with amino acids on either side of the scissile bond in the substrate (D-25, G-27, A-28, G-48, G-49, I-50, L-123, D-125, G-127, P-181, V-182, I-184, R-108, and I-150). In our data set, nine of these sites contain substitutions. Similarly, the

side chains of 13 residues (subsites S2 and S2′) in the protease form the binding subsites that interact with the amino acids that reside two residues from the scissile bond (D-25, G-27, A-28, D-29, D-30, V-32, I-47, G-48, G-49, I-50, R-108, I-84, and I-150). Nine of these residues are substituted (Fig. 1).

Infected persons who have not been treated with protease inhibitors carry variants that appear frequently during selection for inhibitor resistance. Several reports have described mutations that occur in the protease during selection for inhibitor resistance in cell culture (8, 11–13, 24, 31–33, 42) and in vivo during therapy (5). A subset of these mutations have been documented to result in enzymes that are less sensitive to inhibition in vitro (11, 13, 21, 24, 32, 33, 38), and viruses containing some of these substitutions have been demonstrated to be less sensitive to inhibition in cell culture assays. Of interest, other investigators have identified substitutions in the protease that appear regularly during selection for inhibitor resistance in cell culture but do not produce enzymes that are clearly less sensitive to inhibition (8, 11–13, 24, 31–33, 42).

We detected both types of substitutions in a significant proportion of the protease genes sequenced. The R8→Q substitution, which has been demonstrated to increase the K_i for a particular inhibitor (A77003) by 60-fold (13), was found in two of the clones from patient A. Other substitutions that have more modest effects on inhibitor sensitivity, including the V82→I substitution (contained in 54 clones), M46→F (46 clones), and an M46→I substitution (1 clone), were found more commonly. We also identified variants that have appeared in HIV-infected patients during protease inhibitor treatment, including L10→I, T12→I, and L63→P (5). Finally, Partaledis and coworkers have recently reported that an infectious clone carrying an I50→V substitution was less sensitive to inhibition by two hydroxyethylamino sulfonamide protease inhibitors (32). This substitution was found in one of the clones contained in our data set.

Certain combinations of amino acid substitutions occur together either more frequently or less frequently than expected if substitutions occur independently. It was surprising that untreated patients would harbor viral genomes with substitutions in amino acid residues known to be involved either in inhibitor resistance or in direct interactions with the substrate. This was especially unexpected since viruses containing some of these mutations have been demonstrated to grow poorly in cell culture (11, 12a, 18, 24). However, substitutions elsewhere in the protease might compensate for these changes. To evaluate this possibility, clones that contained substitutions at residues known to produce inhibitor-resistant enzymes were evaluated for the presence of mutations elsewhere in the protease coding domain (Table 1). Several of the substitutions appear together quite frequently. When examining the pooled sequences, we found that the M46→F-V82→I combination occurred more frequently than would be expected by chance. Conversely, the V82→I-L63→P and L63→P-M46→F combinations appeared substantially less frequently than would be expected if the appearance of one substitution was independent of the other. This analysis suggests that the double mutation at positions 46 and 82 may play a role in maintaining enzyme function.

We also evaluated the association between these substitutions within subjects. Pooling of sequences obtained from different patients assumes that the entire collection comes from the same population, ignoring patient-to-patient variability. The analysis of pooled sequences can therefore be confounded by the transmission of different initial variants to different patients. Further, rates of evolution of protease sequences may vary from individual to individual, as may the overall viral

TABLE 1. Frequency of covariant mutations

Subject	No. of genes sequenced	Substitutions (%)[a]											
		M46→F	M46→T	L63→P	V82→I	M46→I-L63→P	M46→F-L63→P	M46→F-V82→I	L63→P-V82→I	R8→Q-L63→P	R8→Q-M46→F-V82→I
A	22	4.6		13.6	9.1		4.6	59.1		4.6	4.6
B	20			95.0							
C	22			40.9							
D	18			94.4							
E	18				5.6			55.6			
F	24			91.7		4.2			4.2		
G	18		38.9								
H	23							100.0			
I	24										
J	15			100.0							
K	22			95.5							
L	20			85.0					15.0		

[a] Substitutions as a percentage of total genes sequenced for each subject.

burden in different subjects. Separate analysis for a single patient removes these potentially confounding effects. For a single individual, it is possible to test whether sequence variation is independent only if sufficient variation among the sequences is seen. Two of the subjects, A and E, had variation adequate for this purpose. When the sequences obtained from these subjects were analyzed separately by exact methods, we found that for subject A there was a negative association for both the M46→F-L63→P and L63→P-V82→I combinations ($P = 0.009$ and 0.0002, respectively). There was also a significant three-way association, corresponding to a low probability for the combination M46→F-L63→P-V82→I ($P = 0.0003$). In agreement with our observations concerning the pooled data, there was also a positive association for the M46→F-V82→I combination ($P = 0.025$). Similar results were obtained when the coincident appearance of M46→F and V82→I in subject E was evaluated. In this case, there was a positive association for the two substitutions ($P = 0.0021$).

Infection with HIV-1 is characterized by the generation of a heterogeneous population of quasispecies (1–3, 7, 16, 17, 23, 25–27, 34, 36, 39, 43, 48). Recently, Ho and coworkers and Shaw and coworkers have demonstrated that there is an extremely high rate of viral replication and clearance of infected cells in HIV-infected persons (10, 44). As described by Coffin, the dynamics of population turnover make it likely that viruses containing each single substitution that gives rise to protease inhibitor resistance will be produced every day in an HIV-infected individual (3). The percentage of the virus population composed of these variants is a reflection of their overall relative fitness (2, 3).

Our data set of 246 protease sequences is the largest single collection of such sequences from HIV-1-infected persons. An analysis of the sequence heterogeneity in the protease coding domain indicates that patients who have never received protease inhibitors may harbor proviruses that contain substitutions associated with inhibitor resistance. Since these samples were obtained between 1991 and 1993, before the ready availability of protease inhibitors, it is unlikely that the subjects received these agents surreptitiously. In addition, heterogeneity in the protease may include regions directly involved in substrate-enzyme interactions. Our data indicate that significant variability may be found in the protease and suggest that the enzyme is sufficiently flexible to tolerate substitutions in these critical residues. We also found that certain substitutions appear together more frequently than would be expected by chance alone and that others appear together less frequently than would be expected. The simultaneous appearance of amino acid substitutions at positions 82 and 46 may reflect one substitution compensating for the deleterious effects of the other. Conversely, an underrepresentation of certain combined substitutions seems likely to reflect combinations that are incompatible with adequate enzyme function. Finally, our study was limited to an analysis of mutations in the viral protease. Mutations outside the protease coding domain (i.e., within a cleavage site or sites) may also compensate for mutations within the enzyme.

The structural basis for an interaction between the M46→F-V82→I, L63→P-V82→I, and M46→F-L63→P combinations is difficult to discern given the lack of any direct contact between M-46 and L-63 and the side chains at position 82. Molecular dynamics simulations of the HIV protease suggest that an M46→I substitution influences the protein in a way that is predicted to stabilize the bound conformation of the enzyme (4). Similar results were obtained with the M46→F mutation (8a). It is interesting that these substitutions frequently appear during selection for inhibitor resistance despite the absence of direct interactions between the residues at positions 63 and 46 and the inhibitors. In this regard, it is perhaps noteworthy that the wild-type sequence of the HIV-2 protease also contains the M46→I-V82→I combination. Taken together, these observations strongly imply a synergistic or compensatory role for the M-46–V-8 combination. A proline has been detected at position 63 in many wild-type isolates, and an L63→P substitution has been demonstrated to improve viral replication in the presence of double mutations at positions 82 and 84 (24). The negative association between L63→P and substitutions at residues 46 and 82 may result either from long-range structural perturbations or from the influence of residue 63 on flap dynamics.

Recently, Nájera et al. reported an analysis of sequences of the reverse transcriptase (RT) coding domain from patients who either had or had not received RT inhibitor therapy (29). In patients who had not received antiretroviral therapy, a number of sequences with substitutions predicted to give rise to RT inhibitor-resistant enzymes were detected. This confirms the work of Mohri and coworkers, in which zidovudine-resistant viruses were cultured from RT inhibitor-naive subjects (28). Further, Winslow and coworkers reported protease sequences from three laboratory strains and 21 cultured clinical isolates (45). Their data set of PCR-generated consensus sequences contained few substitutions. In contrast to these reports, all of our data were obtained from individual clones rather than from the consensus sequences obtained by direct sequencing of PCR products. Therefore, our studies were designed to identify both the majority population of protease sequences and those variants which may be present at lower frequencies. This is especially important since it is expected that the application of protease inhibitors would result in the rapid outgrowth of this resistant minority population.

Thus, for the two *pol* gene products currently targeted for therapeutic intervention, the presence of preexisting populations of resistant variants in untreated patients appears to be a general feature of the distribution of viral quasispecies. The generation of protease inhibitor-resistant variants of HIV-1 depends in large part on the ability of the viral protease to tolerate substitutions. Current attempts to subvert viral resistance include combination therapeutic strategies designed to select multiply mutated variants that are relatively attenuated (19, 40). Unfortunately, the ability of the virus to produce compensatory substitutions may provide the protease with flexibility sufficient to vary critical subsites in vivo, thereby retaining enzyme function and viral pathogenicity.

We thank R. Swanstrom for helpful discussions and H. Ly for technical assistance. We also thank J. Collins for assistance with preparation of the figures.

W.J.L. is supported by an award from the NSF-California Alliance for Minority Participation in the Sciences. A.H.K. is supported by NIH awards K11 AI01107 and R29 AI36702. L.C.L. is supported by USPHS grant 5R01 CA 59039-20 and NSF grant DMS-9510516.

REFERENCES

1. **Cichutek, K., H. Merget, S. Norley, R. Linde, W. Kreuz, M. Gahr, and R. Kurth.** 1992. Development of a quasispecies of HIV-1 *in vivo*. Proc. Natl. Acad. Sci. USA **89:**7365–7369.
2. **Coffin, J. M.** 1992. Genetic diversity and evolution of retroviruses. Curr. Top. Microbiol. Immunol. **176:**143–164.
3. **Coffin, J. M.** 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science **267:**483–489.
4. **Collins, J. R., S. K. Burt, and J. W. Erickson.** 1995. Flap opening in HIV-1 protease simulated by activated molecular dynamics. Nat. Struct. Biol. **2:**334–338.
5. **Condra, J. H., W. Schleif, O. M. Blahy, L. J. Gabryelski, D. J. Graham, J. C. Quintero, A. Rhodes, H. L. Robbins, E. Roth, M. Shivaprakash, D. Titus, T. Yang, H. Tepler, K. E. Squires, P. J. Deutsch, and E. A. Emini.** 1995. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature (London) **374:**569–571.

6. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1984. Protein biosynthesis and assembly, p. 513–648. *In* R. A. Weiss, R. A. Teich, H. E. Varmus, and J. M. Coffin (ed.), Molecular biology of tumor viruses: RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **di Marzo Veronese, F., M. S. Reitz, Jr., G. Gupta, M. Robert-Guroff, C. Boyer-Thompson, A. Louie, R. C. Gallo, and P. Lusso.** 1993. Loss of a neutralizing epitope by a spontaneous point mutation in the V3 loop of HIV-1 isolated from an infected laboratory worker. J. Biol. Chem. **268:**25894–25901.
8. **el-Farrash, M. A., M. J. Kuroda, T. Kitazaki, T. Masuda, K. Kato, M. Hatanaka, and S. Harada.** 1994. Generation and characterization of a human immunodeficiency virus type 1 (HIV-1) mutant resistant to an HIV-1 protease inhibitor. J. Virol. **68:**233–239.

8a. **Erickson, J. W.** Unpublished data.

9. **Ferbas, J., A. H. Kaplan, M. A. Hausner, L. E. Hultin, J. L. Matud, Z. Liu, D. Panicali, H. Nerng-Ho, R. Detels, and J. V. Giorgi.** 1995. Immunologic and virologic determinants of survival in chronic human immunodeficiency virus infection. J. Infect. Dis. **172:**329–339.
10. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature (London) **373:**123–126.
11. **Ho, D. D., T. Toyoshima, H. Mo, D. J. Kempf, D. Norbeck, C.-M. Chen, N. E. Wideburg, S. K. Burt, J. W. Erickson, and M. K. Singh.** 1994. Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C_2-symmetric protease inhibitor. J. Virol. **68:**2016–2020.
12. **Jacobsen, H., K. Yasargil, D. L. Winslow, J. C. Craig, A. Krohn, I. B. Duncan, and J. Mous.** 1995. Characterization of human immunodeficiency virus type 1 mutants with decreased sensitivity to proteinase inhibitor Ro 31-8959. Virology **206:**527–534.

12a. **Kaplan, A. H.** Unpublished observation.

13. **Kaplan, A. H., S. F. Michael, R. S. Wehbie, M. F. Knigge, D. A. Paul, L. Everitt, D. J. Kempf, D. W. Norbeck, J. Erickson, and R. Swanstrom.** 1994. Selection of multiple HIV-1 variants with decreased sensitivity to an inhibitor of the viral protease. Proc. Natl. Acad. Sci. USA **91:**5597–5601.
14. **Kaplan, A. H., J. A. Zack, M. Knigge, D. A. Paul, D. J. Kempf, D. W. Norbeck, and R. Swanstrom.** 1993. Partial inhibition of the human immunodeficiency virus type 1 protease results in aberrant virus assembly and the formation of noninfectious particles. J. Virol. **67:**4050–4055.
15. **Kohl, N. E., E. A. Emini, W. A. Schleif, L. J. Davis, J. C. Heimbach, R. A. Dixon, E. M. Scolnick, and I. S. Sigal.** 1988. Active human immunodeficiency virus protease is required for viral infectivity. Proc. Natl. Acad. Sci. USA **85:**4686–4690.
16. **Kozal, M. J., R. W. Shafer, M. A. Winters, D. A. Katzenstein, and T. C. Merigan.** 1993. A mutation in the human immunodeficiency virus reverse transcriptase and decline in CD4 lymphocyte numbers in long-term zidovudine recipients. J. Infect. Dis. **167:**526–532.
17. **Kuiken, C. L., G. Zwart, E. Baan, R. A. Coutinho, J. A. van den Hoek, and J. Goudsmit.** 1993. Increasing antigenic and genetic diversity of the V3 variable domain of the HIV envelope protein in the course of the AIDS epidemic. Proc. Natl. Acad. Sci. USA **90:**9061–9065.
18. **Kuroda, M. J., M. A. El-Farrash, S. Choudhury, and S. Harada.** 1995. Impaired infectivity of HIV-1 after a single point mutation in the *pol* gene to escape the effect of a protease inhibitor *in vitro*. Virology **210:**212–216.
19. **Larder, B. A.** 1992. 3′-Azido-3′-deoxythymidine resistance suppressed by a mutation conferring human immunodeficiency virus type 1 resistance to nonnucleoside reverse transcriptase inhibitors. Antimicrob. Agents Chemother. **36:**2664–2669.
20. **Lazzeroni, L., and K. Lange.** 1995. Markov chains for Monte Carlo tests of genetic equilibrium in multidimensional contingency tables. Stanford University Division of Biostatistics Technical Report no. 178.
21. **Lin, Y., X. Lin, L. Hong, S. Foundling, R. L. Heinrikson, S. Thaisrivongs, W. Leelamanit, D. Raterman, M. Shah, B. M. Dunn, and J. Tang.** 1995. Effect of point mutations on the kinetics and the inhibition of human immunodeficiency virus type 1 protease: relationship to drug resistance. Biochemistry **34:**1143–1152.
22. **Loeb, D. D., R. Swanstrom, L. Everitt, M. Manchester, S. E. Stamper, and C. E. Hutchinson.** 1989. Complete mutagenesis of the HIV-1 protease. Nature (London) **340:**397–400.
23. **Markham, R. B., X. Yu, H. Farzadegan, S. C. Ray, and D. Vlahov.** 1995. Human immunodeficiency virus type 1 env and p17gag sequence variation in polymerase chain reaction-positive, seronegative injection drug users. J. Infect. Dis. **4:**797–804.
24. **Markowitz, M., H. Mo, D. J. Kempf, D. W. Norbeck, T. N. Bhat, J. W. Erickson, and D. D. Ho.** 1995. Selection and analysis of human immunodeficiency virus type 1 variants with increased resistance to ABT-538, a novel protease inhibitor. J. Virol. **69:**701–706.
25. **McCutchan, F. E., P. A. Hegerich, T. P. Brennan, P. Phanuphak, P. Singharaj, A. Jugsudee, P. W. Berman, A. M. Gray, A. K. Fowler, and D. S. Burke.** 1992. Genetic variants of HIV-1 in Thailand. AIDS Res. Hum. Retroviruses **8:**1887–1895.
26. **McNearny, T., Z. Hornickova, R. Markham, A. Birdwell, M. Arens, A. Saah, and L. Ratner.** 1992. Relationship of HIV-1 sequence heterogeneity to stage of disease. Proc. Natl. Acad. Sci. USA **89:**10247–10251.
27. **Milich, L., B. Margolin, and R. Swanstrom.** 1993. V3 loop of the human immunodeficiency virus type 1 Env protein: interpreting sequence variability. J. Virol. **67:**5623–5634.
28. **Mohri, H., M. K. Singh, W. T. Ching, and D. D. Ho.** 1993. Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the blood of treated and untreated patients. Proc. Natl. Acad. Sci. USA **90:**25–29.
29. **Nájera, I., Á. Holguín, M. E. Quiñones-Mateu, M. Á. Muñoz-Fernández, R. Nájera, C. López-Galíndez, and E. Domingo.** 1995. *pol* gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy. J. Virol. **69:**23–31.
30. **Norbeck, D. W., and D. J. Kempf.** 1991. HIV-1 protease inhibitors. Annu. Rep. Med. Chem. **26:**141–150.
31. **Otto, M. J., S. Garber, D. L. Winslow, C. D. Reid, P. Aldrich, P. K. Jadhav, C. E. Patterson, C. N. Hodge, and Y. S. Cheng.** 1993. In vitro isolation and identification of human immunodeficiency virus (HIV) variants with reduced sensitivity to C-2 symmetrical inhibitors of HIV type 1 protease. Proc. Natl. Acad. Sci. USA **90:**7543–7547.
32. **Partaledis, J. A., K. Yamaguchi, M. Tisdale, E. E. Blair, C. Falcione, B. Maschera, R. E. Myers, S. Pazhanisamy, O. Futer, A. B. Cullinan, C. M. Stuver, R. A. Byrne, and D. J. Livingston.** 1995. In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease. J. Virol. **69:**5228–5235.
33. **Patick, A. K., R. Rose, J. Greytok, C. M. Bechtold, M. A. Hermsmeier, P. T. Chen, J. C. Barris, R. Zahler, R. J. Colonno, and P. F. Lin.** 1995. Characterization of a human immunodeficiency virus type 1 variant with reduced sensitivity to an aminodiol protease inhibitor. J. Virol. **69:**2148–2152.
34. **Pedroza-Martins, L., N. Chenciner, and S. Wain-Hobson.** 1992. Complex intra patient sequence variation in the V1 and V2 hypervariable regions of the HIV-1 gp120 envelope sequence. Virology **191:**837–845.
35. **Peng, C., B. K. Ho, T. W. Chang, and N. T. Chang.** 1989. Role of human immunodeficiency virus type 1-specific protease in core maturation and viral infectivity. J. Virol. **63:**2550–2556.
36. **Phillips, R. E., S. Rowland-Jones, D. F. Nixon, F. Gothch, J. P. Edwards, A. O. Ogunlesi, J. G. Elvin, J. A. Rothbard, C. R. M. Bangham, C. R. Rizza, and A. J. McMichael.** 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T-cell recognition. Nature (London) **354:**453–459.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Sardana, V. V., A. J. Schlabach, P. Graham, B. L. Bush, J. H. Condra, J. C. Culberson, L. Gotlib, D. J. Graham, N. E. Kohl, R. L. LaFemina, C. L. Schneider, B. S. Wolanski, J. A. Wolfgang, and E. A. Emini.** 1994. Human immunodeficiency virus type 1 protease inhibitors: evaluation of resistance engendered by amino acid substitutions in the enzyme's substrate binding site. Biochemistry **33:**2004–2010.
39. **Simmonds, P., L. Q. Zhang, F. McOmish, P. Balfe, C. A. Ludlam, and A. J. Leigh Brown.** 1991. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 *env* sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. **65:**6266–6276.
40. **St. Clair, M. H., J. L. Martin, G. Tudor-Williams, M. C. Bach, C. L. Vavro, D. M. King, P. Kellam, S. D. Kemp, and B. A. Larder.** 1991. Resistance to ddI and sensitivity to AZT induced by a mutation in HIV-1 reverse transcriptase. Science **253:**1557–1559.
41. **Tomasselli, A. G., W. J. Howe, T. K. Sawyer, A. Wlodawer, and R. L. Heinrikson.** 1991. The complexities of AIDS: an assessment of the HIV protease as a therapeutic target. Chim. Oggi. **2:**6–27.
42. **Turriziani, O., G. Antonelli, H. Jacobsen, J. Mous, E. Riva, M. Pistello, and F. Dianzani.** 1994. Identification of an amino acid substitution involved in the reduction of sensitivity of HIV-1 to an inhibitor of viral proteinase. Acta Virol. **38:**297–298.
43. **Vartanian, J. P., A. Meyerhans, M. Henry, and S. Wain-Hobson.** 1992. High resolution structure of an HIV-1 quasispecies: identification of non-coding sequences. AIDS **6:**1095–1098.
44. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature (London) **373:**117–122.
45. **Winslow, D. L., S. Stack, R. King, H. Scarnati, A. Bincsik, and M. J. Otto.** 1995. Limited sequence diversity of the HIV type 1 protease gene from clinical isolates and *in vitro* susceptibility to HIV protease inhibitors. AIDS Res. Hum. Retroviruses **11:**107–113.
46. **Wlodawer, A., and J. Erickson.** 1993. Structure-based inhibitors of the HIV-1 protease. Annu. Rev. Biochem. **62:**543–585.
47. **Yang, Y. L., Q. Wang, K. Dorman, and A. H. Kaplan.** Long PCR amplification of heterogeneous HIV-1 templates results in recombination at a high frequency. AIDS Res. Hum. Retroviruses, in press.
48. **Zhang, L. Q., P. MacKenzie, A. Cleland, E. C. Holmes, A. J. Leigh Brown, and P. Simmonds.** 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J. Virol. **67:**3345–3356.