Protease Cleavage Sites in HIV-1 gp120 Recognized by Antigen Processing Enzymes Are Conserved and Located at Receptor Binding Sites[∇][†]

Bin Yu, Dora P. A. J. Fonseca, Sara M. O'Rourke, and Phillip W. Berman*

Department of Biomolecular Engineering, University of California, Santa Cruz, California 95064

The identification of vaccine immunogens able to elicit broadly neutralizing antibodies (bNAbs) is a major goal in HIV vaccine research. Although it has been possible to produce recombinant envelope glycoproteins able to adsorb bNAbs from HIV-positive sera, immunization with these proteins has failed to elicit antibody responses effective against clinical isolates of HIV-1. Thus, the epitopes recognized by bNAbs are present on recombinant proteins, but they are not immunogenic. These results led us to consider the possibility that changes in the pattern of antigen processing might alter the immune response to the envelope glycoprotein to better elicit protective immunity. In these studies, we have defined protease cleavage sites on HIV gp120 recognized by three major human proteases (cathepsins L, S, and D) important for antigen processing and presentation. Remarkably, six of the eight sites identified in gp120 were highly conserved and clustered in regions of the molecule associated with receptor binding and/or the binding of neutralizing antibodies. These results suggested that HIV may have evolved to take advantage of major histocompatibility complex (MHC) class II antigen processing enzymes in order to evade or direct the antiviral immune response.

A major goal of HIV vaccine development is the development of immunogens that elicit protective antiviral antibody and cellular immune responses. However, after more than 25 years of research, vaccine immunogens able to elicit protective immunity in humans have yet to be described (11, 31). Although it has been possible to produce recombinant envelope proteins (gp120 and gp140) with many of the features of native virus proteins (e.g., complex glycosylation and the ability to bind CD4, chemokine receptors, and neutralizing antibodies), these antigens have not been able to elicit broadly neutralizing antibodies (bNAbs) or protective immune responses when used as immunogens (11, 32, 43, 50, 56, 74, 79). The fact that recombinant proteins can adsorb virus bNAbs from HIV-1-positive sera (59, 91) indicates that many recombinant envelope proteins are correctly folded but that the epitopes recognized by bNAbs are simply not immunogenic. Over the last decade, several different approaches have been employed to create immunogens able to elicit broadly neutralizing antibodies. These strategies have included efforts to duplicate and/or stabilize the oligomeric structure of HIV envelope proteins (5, 26, 87), the creation of minimal antigenic structures lacking epitopes that conceal important neutralizing sites (27, 46, 70, 89), and prime/boost strategies combining protein immunization with DNA immunization or infection with recombinant viruses in order to stimulate the endogenous synthesis and presentation of HIV immunogens (15, 29, 30, 83). However,

none of these approaches has resulted in a clinically significant improvement in antiviral immunity or HIV vaccine efficacy. Efforts to elicit protective cellular immune responses (e.g., cytotoxic lymphocytes) by use of recombinant virus vaccines have likewise been disappointing (10, 61). In fact, such vaccines may have promoted HIV infection rather than inhibiting it (22, 23).

In the present study, we describe the first steps in a new approach to reengineering the immunogenicity of HIV envelope proteins in order to improve the potency and specificity of humoral and cellular immune responses. The approach is based on defining the determinants of antigen processing and presentation of HIV envelope glycoproteins. Both humoral and cellular immune responses depend on proteolytic degradation of protein antigens prior to antigen presentation, mediated by professional antigen-presenting cells (APCs) such as macrophages, dendritic cells, and B cells (97). Normally, proteins of intracellular origin are processed by the proteasome, a 14- to 17-subunit protein complex located in the cytosol. Proteins of extracellular origin are processed in lysosomes or late endosomes of APCs. The resulting peptide epitopes are then loaded into major histocompatibility complex (MHC) class I or class II molecules and presented on the surfaces of APCs to CD8 or CD4 T cells. Within the endosomes and lysosomes of APCs, there are cathepsins, acid thiol reductase, and aspartyl endopeptidase. The enzymes perform two activities: degrading endocytosed protein antigens to liberate peptides for MHC class II binding (99) and removing the invariant chain chaperone (6, 94). Although all cathepsins can liberate epitopes from a diverse range of antigens (16), only cathepsins S and L have nonredundant roles in antigen processing *in vivo* (reviewed by Hsing and Rudensky [45]). Cathepsin L is expressed in thymic cortical epithelial cells but not in B cells or dendritic cells, while cathepsin S is found in all three types of APCs. Unlike

* Corresponding author. Mailing address: Department of Biomolecular Engineering, Baskin School of Engineering, University of California, Santa Cruz, 1156 High Street, MS-SOE2, Santa Cruz, CA 95064. Phone: (831) 459-3529. Fax: (831) 459-1970. E-mail: pwb@soe.ucsc.edu.

cathepsins L and S, which are cysteine proteases and active at neutral pH, cathepsin D is an aspartic protease, is active at acidic pH, and participates in proteolysis and antigen presentation in connection with MHC class I and class II antigen presentation pathways established for CD4 and CD8 T cells. In considering the use of envelope proteins as potential vaccines, the route of immunization, formulation (e.g., adjuvants), protein folding, disulfide bonding, and glycosylation pattern all determine which peptides are available for MHC-restricted presentation.

Previous studies provided evidence that gp120 was sensitive to digestion by cathepsins B, D, and L, but the specific cleavage sites were not defined (18). In the present study, we (i) describe the locations of eight protease cleavage sites on HIV-1 gp120 recognized by cathepsins L, S, and D, involved in antigen processing; (ii) determine the extent to which they are conserved; and (iii) evaluate the effect of cathepsin cleavage on the binding of gp120 to CD4-IgG and neutralizing antibodies. The results obtained provide new insights into the basis of envelope immunogenicity that may prove to be useful in the development of HIV vaccine antigens.

MATERIALS AND METHODS

Proteins, enzymes, and enzyme inhibitors. Recombinant gp120 from the MN strain of HIV-1 (MN-rgp120) was produced in Chinese hamster ovary (CHO) cells by Genentech, Inc. (South San Francisco, CA). MN-rgp120 was a major component of the candidate HIV vaccine AIDSVAX B/B (32). This protein differs from native gp120 in that the first 12 amino acids from the mature form of gp120 were deleted and replaced with a 30-amino-acid flag epitope from the amino terminus of herpes simplex virus (HSV) glycoprotein D. Purified human cathepsins L, S, and D as well as the cathepsin L and D inhibitors N-acetyl-Leu-Leu-methional calpain inhibitor II (ALLM) and pepstatin A were obtained from Biomol (Philadelphia, PA).

MAbs. The broadly neutralizing, CD4-blocking monoclonal antibody (MAb) b12 (12, 70) was obtained from Polymun Scientific (Vienna, Austria). The virus entry inhibitor CD4-IgG (13) and MAbs to the MN and IIIB strains of HIV, reactive with the V3 domain (1026), the C4 domain (13H8), and the V2 domain (6E10 and 1088), were obtained from Genentech, Inc. (South San Francisco, CA), and have been described previously (66, 67). The neutralizing MAb 447-52D (24) was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH, from Susan Zolla-Pazner. Polyclonal antibody D7324 was purchased from Aalto Bio Reagents Ltd. (Dublin, Ireland). Horseradish peroxidase (HRP)-labeled goat anti-human IgG and goat anti-mouse IgG+M were obtained from American Qualex Antibodies (San Clemente, CA).

Cathepsin digestions for cleavage site analysis. Fifty micrograms of MN-rgp120 in 25 μl of 100 mM sodium acetate, pH 5.5, digestion buffer was mixed with 0.5 μg cathepsin L (protease-to-protein ratio, 1:100). The reaction mix was incubated at 37°C. Aliquots of 3 μl were taken at 15 min, 30 min, 60 min, 2 h, 4 h, and 7 h, and the digestion was stopped by rapid cooling in liquid nitrogen. An additional 3-μl aliquot was taken after overnight incubation at room temperature. The aliquots from cathepsin L digestion were mixed with 3 μl of 3× reducing polyacrylamide gel electrophoresis (PAGE) sample buffer (5% SDS, 5% β-mercaptoethanol, 40% glycerol, and 200 mM Tris, pH 6.8) and boiled for 2 min. The collected samples were run in two 4 to 12% Bis-Tris precast gels (Invitrogen, Carlsbad, CA). Digested fragments were visualized either by direct Coomassie blue staining or on immunoblots after electrophoresis onto a polyvinylidene difluoride (PVDF) membrane (Millipore Immobilon PSQ). For sequencing of peptides on PVDF membranes, bands were cut out and transferred to the Molecular Structure Facility at the University of California, Davis, for N-terminal protein sequencing by Edman degradation. The same experimental procedure was carried out for cathepsin S and D digestion, except for the digestion buffer, which was 50 mM sodium phosphate, pH 6.5, with 50 mM sodium chloride for cathepsin S and 100 mM sodium acetate, pH 3.3, for cathepsin D.

Cathepsin digestions for ELISA. To prepare cathepsin L-digested MN-rgp120 for enzyme-linked immunosorbent assay (ELISA), 25 μg of MN-rgp120 in 50 μl

of 100 mM sodium acetate, pH 5.5, digestion buffer was mixed with 1 μg cathepsin L (protease-to-protein ratio, 1:25) at 37°C for overnight incubation, followed by 1 μl ALLM (25 mg/ml in dimethyl sulfoxide [DMSO]) solution to stop the digestion reaction. For cathepsin D-treated MN-rgp120, 25 μg of MN-rgp120 was mixed with 1 μg of cathepsin D in 50 μl buffer (100 mM sodium acetate, pH 3.3) at 37°C for 1 h. Pepstatin A (1 μl at 25 mg/ml in DMSO) solution was added to stop cathepsin D activity.

Detection of monoclonal antibodies and CD4-IgG binding to cathepsin-digested MN-rgp120. Wells of microtiter plates (Immulon II; Thermo-Fisher Scientific) were coated with 100 μl of polyclonal antibody D7324 solution (2 μg/ml in phosphate-buffered saline [PBS]) overnight at 4°C. The wells were blocked with 200 μl of blocking buffer (1% bovine serum albumin [BSA] in PBS) and incubated at 37°C for 1 h. After rinsing of the plates with washing buffer (0.05% Tween 20 in PBS), 100 μl of cathepsin L-treated, cathepsin D-treated, or undigested MN-rgp120 solution was added to each well (2 μg/ml in blocking buffer) and incubated for 1 h at 37°C. After washing of the plates, MAbs and CD4-IgG were added and fivefold serial dilutions were carried out, starting with 25 μg/ml of MAb b12 and 5 μg/ml of all other MAbs and CD4-IgG, and incubated for 1 h at 37°C. After washing of the plates three times with washing buffer, 100 μl of HRP-labeled goat anti-human IgG or goat anti-mouse IgG+M solution (diluted 1:10,000 in blocking buffer) was added and incubated for 1 h at 37°C. Finally, after plate washing, 100 μl of 0.4-mg/ml o-phenylenediamine dihydrochloride (Sigma-Aldrich Chemicals, St. Louis, MO) solution was added and incubated at room temperature for 10 min, followed by 100 μl of 3 M sulfuric acid to stop the reaction. The absorbance was measured by a SpectraMax 190 plate reader (Molecular Devices) at 490 nm.

Prediction of cleavage sites by computational methods. Envelope glycoprotein sequences were obtained from various sources (described below) and aligned using MAFFT (52). The sequence for gp120 from the MN strain of HIV-1 used in these studies, MN_{GNE}, differs from the sequence reported by Gurgo et al. (41) and has been published previously (71). To determine the locations of predicted cathepsin cleavage sites in MN-rgp120, we used the PoPs program developed by Boyd et al. (8), cleavage specificity algorithms for cathepsins L, S, and B generated by Choe et al. (19), and the cathepsin D recognition sequence of Scarborough et al. (85).

Conservation of protease cleavage sites. Three data sets were used to investigate the sequence conservation of cathepsin cleavage sites. The VAX004 data set (32) was obtained from the Global Solutions for Infectious Diseases (GSID) HIV data browser (http://www.gsid.org/), which includes 1,047 clade B envelope glycoprotein sequences from 349 individuals with recent HIV infections. A data set of acute and recent clade B infections containing 2,908 envelope glycoprotein sequences from 102 infected individuals was obtained from the studies of Keele et al. (53). Finally, a listing of clade-specific reference sequences, as well as a data set containing 1,766 envelope glycoprotein sequences from isolates collected worldwide at various undefined times after HIV infection, was obtained from the Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov/). The sequences from all three databases were aligned using MAFFT (52).

RESULTS

Computational methods to locate protease cleavage sites. Cathepsins L, S, and D are known to play an important role in antigen processing and presentation (28, 44, 45, 94). In initial studies, we used computational methods (see Materials and Methods) to determine whether gp120 was likely to possess cleavage sites recognized by cathepsins known to be important for antigen processing. For these studies, we examined sequences with the prediction algorithm (8) set for maximum stringency. The results of these studies (see Table S1 in the supplemental material) suggested that MN-rgp120 was likely to possess multiple cathepsin cleavage sites. However, because cathepsin cleavage sites are difficult to predict, we reasoned that actual protease digestion studies would be required to reliably identify the number and locations of these sites.

Mapping cathepsin L cleavage sites. Initially, we examined the sensitivity of MN-rgp120 to digestion by cathepsin L. A time course experiment is shown in Fig. 1A. We found that cathepsin L digestion resulted in six proteolytic fragments.
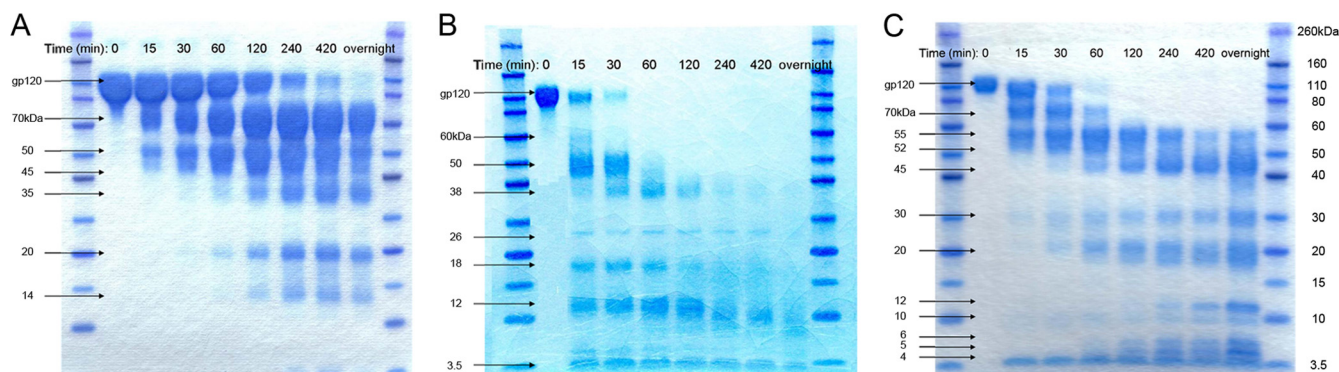
FIG. 1. Cathepsin digestion of MN-rgp120. MN-rgp120 was digested with either cathepsin L (A), cathepsin S (B), or cathepsin D (C). At the indicated times, samples were removed, the digestion was stopped by the addition of liquid nitrogen, and samples were prepared for SDS-PAGE analysis. Bands were visualized by Coomassie blue staining. The mobilities of the fragments identified are shown to the left of each panel. The mobility values of prestained molecular mass markers are shown to the right of panel C.

Because of their size, these fragments could not be analyzed by mass spectrometry but, rather, required analysis by Edman sequence degradation. The sizes, locations, and experimentally determined N-terminal sequences of the peptides isolated are shown in Table 1. The scissile peptide bond represents that joining the P1 and P1′ residues, according to the nomenclature of Schechter and Berger (86). A listing of flanking residues likely to include the enzyme recognition residues and ranging from P4 to P4′ is provided in Table S2 in the supplemental material. We found that digestion with cathepsin L resulted in a 70-kDa fragment and a 50-kDa fragment that appeared within 15 min of treatment. Edman degradation showed that the first five amino acids in the N terminus of the 50-kDa

fragment are GTIRQ, which revealed that the cleavage site is located between the K_{327} and G_{328} residues in the V3 domain. The N terminus of the 70-kDa fragment is derived from cleavage between the A_3-L_4 bond in the glycoprotein D flag epitope at the N terminus of MN-rgp120, resulting in the L_4-A_5-D_6 N-terminal sequence. The kinetics of the appearance of these two fragments indicate that MN-rgp120 was first attacked at the V3 domain cleavage site, K_{327}-G_{328}, resulting in the 70-kDa fragment and the 50-kDa fragment. The 50-kDa fragment was subsequently degraded with longer digestion times to yield additional fragments (Fig. 1A and 2). Edman degradation confirmed that the resulting 45-kDa and 35-kDa fragments were originally from the 50-kDa fragment, because both included

TABLE 1. Identification of cathepsin cleavage sites in MN-rgp120 by N-terminal sequencing

Protease	Fragment size (kDa)	Cleavage site	N-terminal sequence	Domain	Associated function[a]
Cathepsin L	70	A_3-L_4	LAD	HSV-1 gD	Flag
	50	K_{327}-G_{328}	GTIRQ	V3	3
	45	K_{327}-G_{328}	GTIRQ	V3	3
	35	K_{327}-G_{328}	GTIRQ	V3	3
	20	G_{431}-K_{432}	KAMY	C4	1, 2, 4
	14	Y_{435}-A_{436}	APPI	C4	2, 4
Cathepsin S	60	Q_{208}-A_{209}	ACPK	C2	1, 2
	50	S_{261}-T_{262}	TQLL	C2	3
	38	T_{322}-T_{323}	TKNI	V3	3
	18	Y_{435}-A_{436}	APPI	C4	2, 4
	12	Y_{435}-A_{436}	APPI	C4	2, 4
	3.5	NA	NA	NA	NA
Cathepsin D	70	L_{181}-Y_{182}	YKLD	V2	3, 5
	55	E_{274}-V_{275}	VVIRS	C2	NA
	52	E_{274}-V_{275}	VVIRS	C2	NA
	45	NA	NA	NA	NA
	30	E_{274}-V_{275}	VVIRS	C2	NA
	20	L_{181}-Y_{182}	YKLD	V2	4, 5
	12	E_{274}-V_{275}	VVIRS	C2	NA
	10	NA	NA	NA	NA
	6	NA	NA	NA	NA
	5	NA	KYAL	HSV-1 gD	Flag
	4	NA	KYAL	HSV-1 gD	Flag

[a] 1, CD4 binding site; 2, chemokine receptor binding site; 3, V3 domain neutralizing antibody binding site; 4, b12 antibody binding site; 5, α4β7 binding site; NA, data not available; and Flag, sequence from herpes simplex virus glycoprotein D used as a flag epitope to facilitate purification.
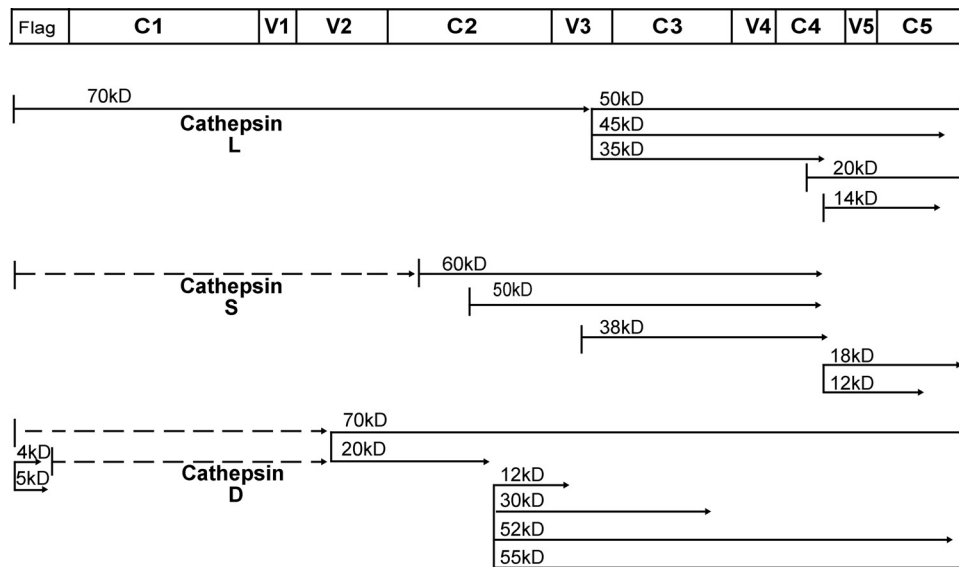
FIG. 2. Diagram of MN-rgp120 fragments generated by cathepsin digestion. MN-rgp120 was digested with cathepsin L, S, or D, and the resulting fragments were purified and analyzed by Edman sequence degradation. Solid lines indicate peptides that were resolved by polyacrylamide gel electrophoresis and characterized by N-terminal sequence analysis. Dashed lines indicate the locations of peptides that were deduced from mobility and sequence analyses but were not recovered.

the same N-terminal sequence, GTIRQ, as that of the 50-kDa fragment. Although the C-terminal sequences of the 50-kDa, 45-kDa, and 35-kDa fragments are not known, at least two cathepsin L cleavage sites in the 50-kDa fragment are indicated, which result in the 45-kDa and 35-kDa fragments. The N-terminal amino acid sequences of the resulting 20-kDa fragment and 14-kDa fragment prove that there are two cathepsin cleavage sites within the C4 domain. The first four amino acids of the N termini of the 20-kDa fragment and the 14-kDa fragment are KAMY and APPI, respectively. Thus, two cathepsin L cleavage sites were identified, located at the G_{431}-K_{432} and Y_{435}-A_{436} bonds. However, because the molecular mass difference between the 20-kDa and 14-kDa fragments is about 6 kDa, while the N-terminal sequence difference between the 20-kDa and 14-kDa fragments is only four amino acids, we deduced that another cathepsin L cleavage site must be present between A_{436} and the C terminus of MN-rgp120.

Mapping cathepsin S cleavage sites. We next examined the ability of cathepsin S to digest gp120, using the same methods. The results of a time course experiment are shown in Fig. 1B. It can be seen that six degradation products were visible on SDS-PAGE gels. The sizes of the peptides isolated, the N-terminal sequences, and the locations within gp120 are shown in Table 1. Compared to cathepsin L digestion, the kinetics of cathepsin S digestion were much more rapid and indicated significantly increased sensitivity to cathepsin S. Six major digestion fragments appeared on the SDS-PAGE gel within 15 min of cathepsin S digestion, indicating greater exposure or accessibility of cathepsin S cleavage sites than of cathepsin L cleavage sites. Because of the rapid digestion by cathepsin S, it was not possible to determine whether there was a kinetically distinct, ordered degradation of gp120, as seems to be the case with cathepsin L. Rather, cathepsin S appears to follow a different digestion pathway, where the protease generates mul-

tiple fragments in a very short time frame. Analysis of six cathepsin S digestion fragments (60 kDa, 50 kDa, 38 kDa, 26 kDa, 18 kDa, and 12 kDa) identified four distinct cathepsin S cleavage sites (Fig. 2). The 26-kDa fragment was derived from cathepsin S. Two of the fragments were located in the C2 domain of gp120, between Q_{208}-A_{209} (60-kDa fragment) and S_{261}-T_{262} (50-kDa fragment). The third cathepsin S cleavage site occurred in the V3 domain and involved the bond joining T_{322} and T_{323} (38-kDa fragment). Finally, an additional cleavage site was located in the C4 domain and occurred at Y_{435}-A_{436} (18-kDa and 12-kDa fragments), which is also a cathepsin L cleavage site. Fragments located N-terminal to the C2 domain were not recovered, suggesting that this region of the molecule contains multiple, as-yet-undefined cathepsin S cleavage sites. It is possible that some of these yield 3.5-kDa fragments, since the final 3.5-kDa band on the PAGE gels appeared to be heterogeneous, with multiple fragments migrating at the same position.

Mapping cathepsin D cleavage sites. A complicated digestion pattern was observed in the digestion of MN-rgp120 with cathepsin D (Fig. 1C). Eleven digestion fragments were visualized on the SDS-PAGE gel, but only eight fragments were able to be characterized by Edman degradation, due to heterogeneity in bands and/or their low abundance. Four fragments (55 kDa, 52 kDa, 30 kDa, and 12 kDa) share the N-terminal sequence VVIRS, which is located in the C2 domain, suggesting the cleavage site E_{274}-V_{275}. Based on differences in molecular masses, we deduced that additional cathepsin D cleavage sites occur in the V3, C3, V4, and C4 domains (Fig. 2). N-terminal sequencing of the 20-kDa and 70-kDa fragments indicated another cathepsin D cleavage site, in the V2 domain, at the bond between residues L_{181} and Y_{182}. The 4-kDa and 5-kDa fragments were from the N terminus and shared the KYAL sequence derived from the HSV-1 gD flag epitope.
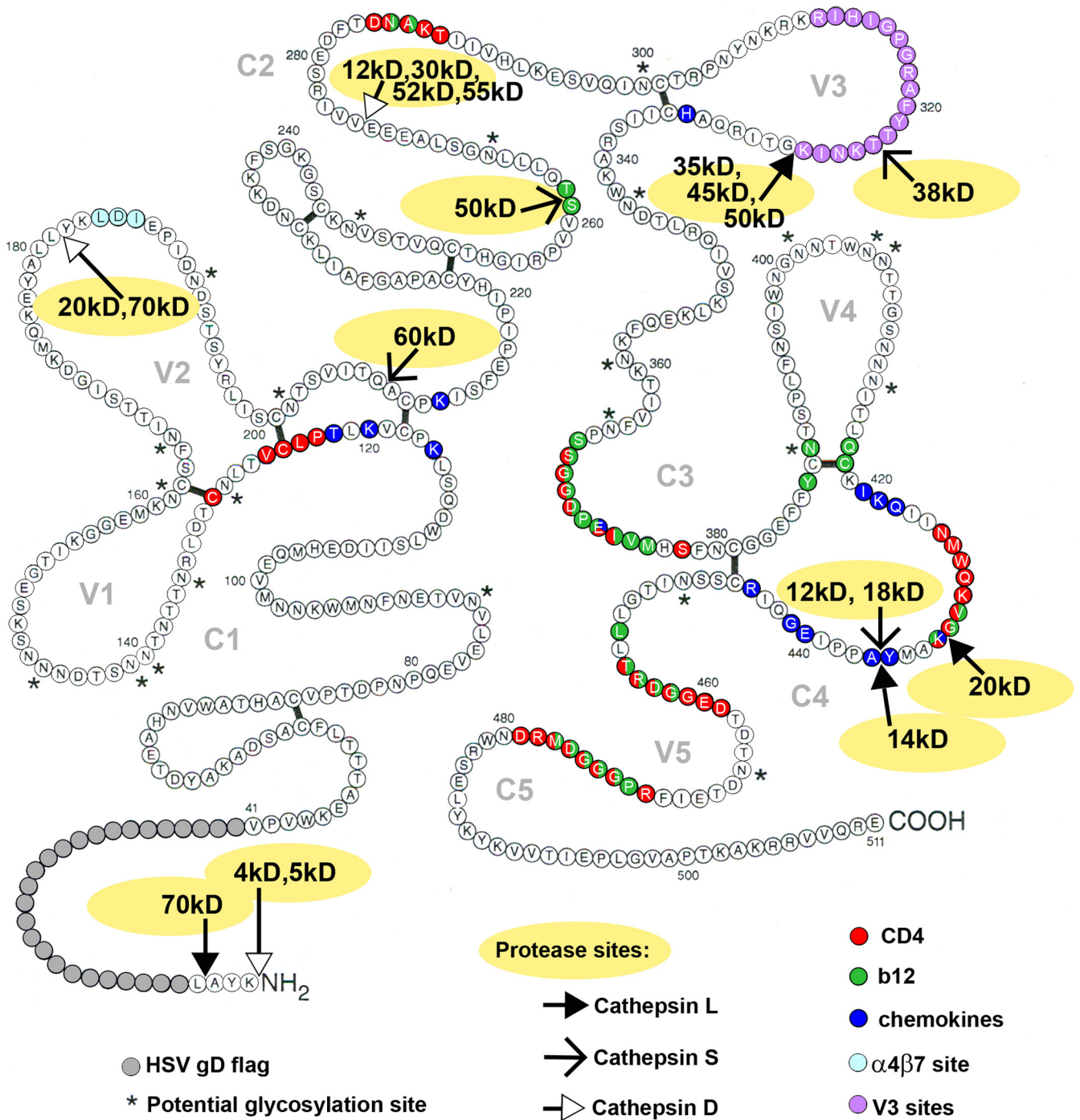
FIG. 3. Locations of cathepsin L, S, and D cleavage sites on MN-gp120 disulfide-bonded schematic. Cathepsin L, S, and D cleavage sites in MN-rgp120 were identified by Edman sequence degradation and located on the disulfide-bonded structure of gp120 determined by Leonard et al. (58). Cathepsin L sites are indicated by closed arrows, cathepsin S sites are indicated by thin-line arrows, and cathepsin D sites are indicated by open arrows. The locations of amino acid residues reported to be important for the binding of CD4, chemokine receptors, and the α4β7 integrin are indicated by residues shaded red, dark blue, and light blue, respectively. The locations of amino acids known to be important for binding of the broadly neutralizing antibody b12 and neutralizing antibodies to the V3 domain are indicated by amino acids shaded green and purple, respectively. Residues with two or more colors indicate amino acids involved in the binding of two or more receptors or neutralizing monoclonal antibodies. The figure was created based on results from Kwong et al. (55), Zhou et al. (105), Rizzuto et al. (79), Decker et al. (25), and Arthos et al. (2). The numbering provided is based on the sequence of gp120 from the MN$_{GNE}$ isolate of HIV (71).
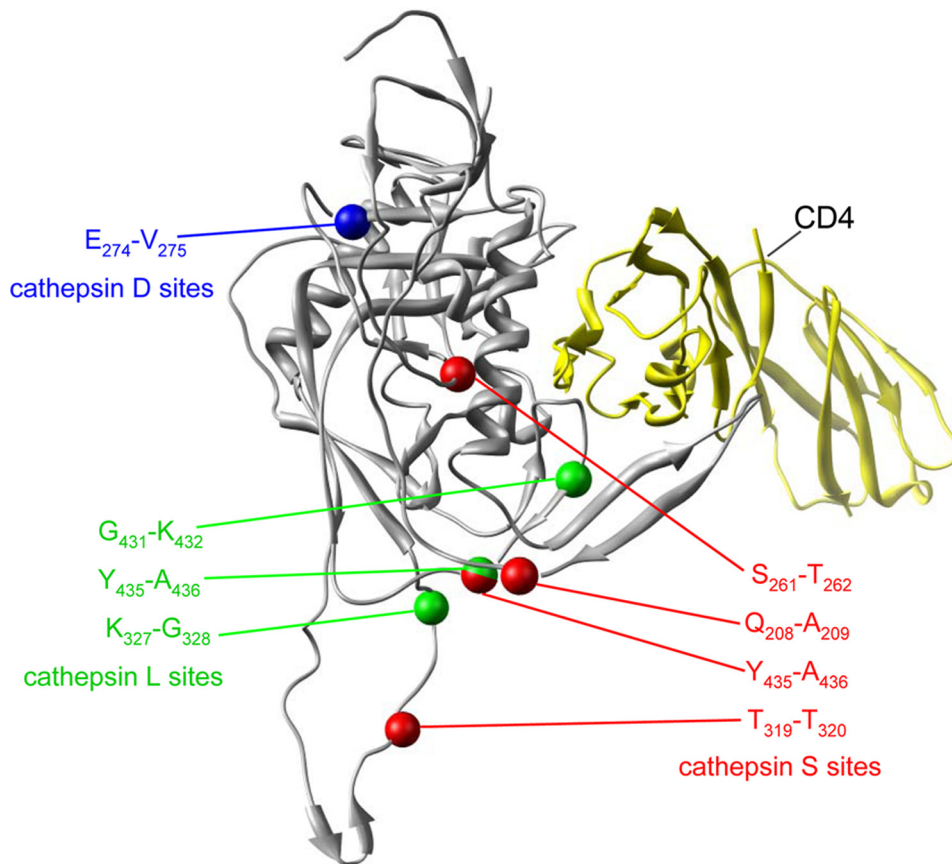
FIG. 4. Locations of cathepsin L, S, and D cleavage sites on three-dimensional structure of gp120 bound to CD4. The locations of cathepsin cleavage sites were located on the three-dimensional structure of gp120 based on the structure of a gp120 fragment complexed with CD4, described by Huang et al. (47). Cathepsin L sites are indicated in green, cathepsin S sites are in red, and cathepsin D sites are in blue. The structure of CD4 is shown in yellow. Numbering is based on the sequence of MN-rgp120.

Locations of cathepsin cleavage sites on the structure of gp120. The locations of cathepsin cleavage sites relative to the disulfide bonds and conserved (C) and variable (V) regions of gp120 were mapped onto the two-dimensional structure of Leonard et al. (58) (Fig. 3) and the three-dimensional structure of Huang et al. (48) (Fig. 4). Also shown in Fig. 3 are the locations of amino acids reported to be involved in receptor binding and the binding of neutralizing antibodies. In total, nine cathepsin cleavage sites were identified, with one in the N-terminal flag sequence and eight in gp120. One cleavage site was found in the V2 domain, three in the C2 domain, two in the V3 domain, and two in the C4 domain.

It was clear from these studies that the cathepsin cleavage sites are not randomly distributed. Remarkably, they appeared to cluster in regions of functional significance, often in close proximity to the binding sites for the CD4 and chemokine coreceptors and/or epitopes recognized by neutralizing antibodies (Table 1). For example, the cathepsin S cleavage sites in the C2 (Q_{208}-A_{209}) and C4 (Y_{435}-A_{436}) domains and the cathepsin L sites at G_{431}-K_{432} and Y_{435}-A_{436} are located in close proximity in the three-dimensional structure of gp120. The K_{432} residue and the Y_{435} residue are reported to be contact residues for chemokine receptor binding, and the Q_{208}-A_{209} cleavage site is three amino acids away from K_{212}, which is also

thought to be a chemokine receptor contact residue (25, 78, 79). Additionally, the G_{431} residue is located within a string of eight amino acids (425 to 432) known to be contact residues for CD4 binding (55). G_{431} and K_{432}, along with V_{429}, are known to be contacts for CD4, chemokine receptors, and the broadly neutralizing b12 MAb (105). Two additional cathepsin cleavage sites occur in the C2 domain. Of these, position T_{262} in the S_{261}-T_{262} cathepsin L cleavage site is known to be a contact residue for the broadly neutralizing b12 MAb (105), whereas the cathepsin D cleavage site (E_{274}-V_{275}) was the only cathepsin cleavage site that was not part of or adjacent to a receptor or neutralizing antibody binding site. Two cathepsin cleavage sites were identified in the V3 domain. The V3 domain is thought to be an important determinant of chemokine receptor usage (21, 101) and is known to possess epitopes recognized by a variety of neutralizing antibodies. A cathepsin S site (T_{322}-T_{323}) is located one amino acid away from the crown of the V3 loop containing the GPGRAF sequence, important for the binding of multiple neutralizing antibodies (24, 66, 81). The cathepsin L cleavage site at K_{327}-G_{328} is four amino acids from the cathepsin S site between the stem and the crown of the V3 loop. Finally, a single cathepsin D site involving residues L_{181} and Y_{182} is located in the V2 domain. The V2 domain is known to possess multiple epitopes for neutralizing

TABLE 2. Conservation of cathepsin cleavage sites in HIV sequence data sets[a]

Domain	Cathepsin	Location of cleavage site	Residues in MN sequence	Polymorphism (prevalence [%]) in HIV cohort		
				VAX004 ($n = 1,047$)	Keele ($n = 2,908$)	Los Alamos ($n = 1,766$)
V2	D	181–182	LY	FY (81.4)	FY (81.0)	FY (92.2)
C2	S	208–209	QA	QA (97.6)	QA (96.8)	QA (97.3)
C2	S	261–262	ST	ST (97.4)	ST (99.3)	ST (97.9)
C2	D	274–275	EV	EV (61.3)	EV (60.6)	EI (39.8)
V3	S	322–323	TT	AT (59.4)	AT (49.5)	AT (58.0)
V3	L	327–328	KG	IG (87.1)	IG (91.6)	IG (83.2)
C4	L	431–432	GK	GK (91.5)	GK (96.3)	GK (36.6)
C4	S, L	435–436	YA	YA (98.9)	YA (97.7)	YA (97.7)

[a] The VAX004 data set for clade B viruses from the United States was obtained from the GSID HIV Sequence Database (http://www.gsid.org/), a data set of clade B viruses from acute infections was obtained from the work of Keele et al. (53), and a data set of worldwide isolates of HIV was obtained from the Los Alamos HIV Sequence Database (http://hiv.lanl.gov).

antibodies (62, 72) and contains the newly described receptor binding site for the α4β7 integrin (2). Interestingly, the L_{181}-Y_{182} cleavage site is located one amino acid away from the LDI/LDV recognition sequence required for α4β7 binding to gp120.

Conservation of cathepsin cleavage sites. An important question in these studies was to determine whether any of the cathepsin protease sites were conserved. In view of the high degree of sequence variation within HIV and the fact that the envelope protein is the most variable of all of the HIV proteins, it was uncertain whether any of the sites would be conserved. In initial studies, we aligned the MN-rgp120 and HXB2 gp120 sequences with 12 reference sequences: 2 from each of four major group clades, A, C, D, and E (crf A/E), plus 2 from the chimpanzee isolate HIV_{cpz} and 2 simian immunodeficiency virus (SIV) sequences (SIV_{Mac251} and SIV_{Mac239}). The results of this analysis are shown in Fig. S1 in the supplemental material, where both the locations and conservation of the sites recognized by cathepsins L, S, and D can be seen along with the locations of predicted cathepsin cleavage sites. This analysis of the residues occurring at the P1 and P1′ positions showed a high level of conservation at six of the eight cathepsin cleavage sites. Remarkably, two sites, including one cathepsin S site, S_{261}-T_{262}, in the C2 domain, and one cathepsin L site, G_{431}-K_{432}, in the C4 domain, were conserved in the reference strains of the major group HIV clades, the HIV_{cpz} strains, and the SIV strains. A high level of conservation (~98%) was also noted at the Q_{208}-A_{209} cleavage site in the C2 domain and the Y_{435}-A_{436} site in the C4 domain. A somewhat lower (81 to 92%) level of conservation was also noted at the L_{181}-Y_{182} site in the V2 domain; however, in this case, the MN strain is unusual in that F is replaced with L at position 181. Comparison of sequence alignments showed that in most cases the cathepsin L site I_{327}-G_{328} is conserved; however, MN-rgp120 is unusual in this respect, with K rather than I at position 327. The highly conserved nature of these sites suggests that they are important for virus function or survival and thus have been preserved by positive selection across species and time.

To further explore the conservation of cathepsin cleavage sites, we examined three independent HIV sequence data sets. One data set (GSID HIV Sequence Database) included 1,047 gp120 sequences from 349 individuals with new and recent HIV infections (less than 6 months postinfection) from differ-

ent cities throughout North America (32). A second data set was obtained from the studies of Keele et al. (53) and consisted of 2,908 sequences from 102 new and acute infections collected in the United States. The third HIV data set examined was the Los Alamos HIV Sequence Database, comprising 1,766 gp120 sequences collected from worldwide isolates that include sequences from the 1980s through the present, mostly from chronic HIV infections. The results of this analysis are presented in Table 2. We found very high levels of conservation (i.e., >96%) for the Q_{208}-A_{209} and S_{261}-T_{262} cathepsin cleavage sites, located in the C2 domain, and for the G_{431}-K_{432} and Y_{435}-A_{436} cleavage sites, located in the C4 domain of gp120. In the case of the 431-432 cleavage site, a significant discrepancy was noted between the Los Alamos data set and the VAX004 and Keele data sets. Further analysis indicated that this result could be attributed to clade-specific polymorphism, as clade B viruses typically possessed K at position 432, while other clades typically possessed R or Q at this position. The L_{181}-Y_{182} cleavage site in the V2 domain was less conserved (i.e., >80%); however, as noted earlier, the sequence of HIV_{MN} was unusual in that F was replaced with L at position 181. However, this replacement probably preserved the cleavage site, since leucine or aromatic residues are known to be favored by cathepsin D (96).

Effect of cathepsin cleavage on binding of CD4-IgG and neutralizing antibodies. Based on the locations of cathepsin cleavage sites at or near receptor binding sites and epitopes recognized by neutralizing antibodies, it was of interest to determine whether cathepsin cleavage actually affected the binding of antibodies to these sites. The binding of monoclonal antibodies to cathepsin-treated and untreated MN-rgp120 was investigated by ELISA (Fig. 5). One concern in performing this assay was the possibility that enzyme cleavage would release small peptide fragments that would not be captured on the microtiter plate. Examination of the protease cleavage sites in relation to the disulfide structure showed that proteolysis of the peptide backbone would not necessarily release multiple peptide fragments, since most would remain associated by virtue of disulfide bonds. Thus, treatment with cathepsin L should release only a small, 4-amino-acid peptide, K_{432} to Y_{435}, from the C4 domain. Treatment with cathepsin D might split the molecule into two large fragments by virtue of the cleavage site located at position 274 in the C2 domain and might also result
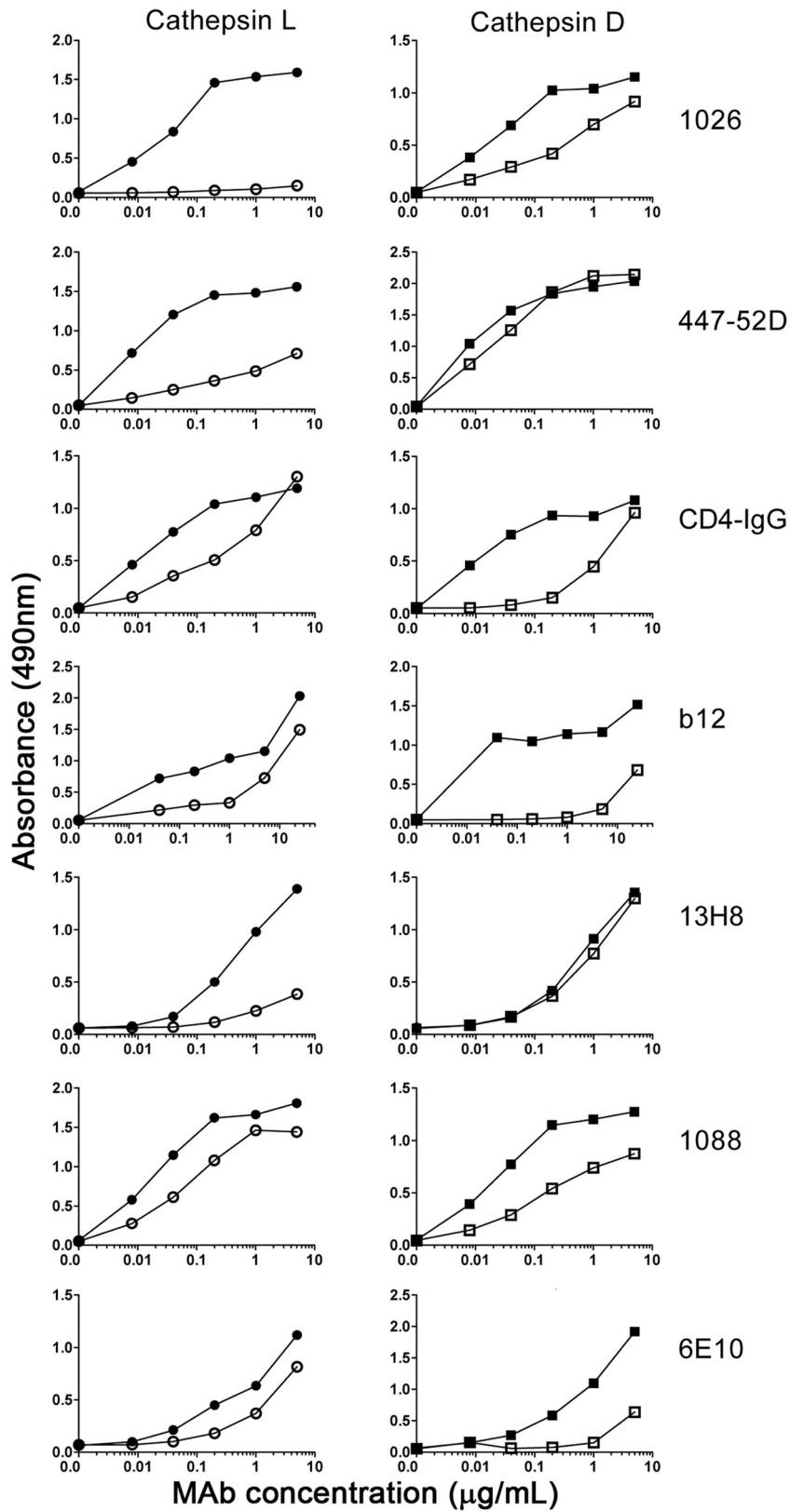
FIG. 5. Antibody binding to cathepsin L- and D-treated gp120. Purified MN-rgp120 was treated with either cathepsin L or cathepsin D and captured on the surfaces of microtiter plates coated with a polyclonal antibody, D7324, directed to the C terminus of gp120. Monoclonal antibodies were incubated with the gp120-coated microtiter plates, and binding was determined by ELISA. Closed symbols indicate binding to untreated gp120, open circles represent binding to cathepsin L-treated gp120, and open squares indicate binding to cathepsin D-treated gp120.

in the release of an undefined, 4- to 5-kDa fragment from the C1 domain. Treatment with cathepsin S should have the largest effect and should result in the loss of the C1, V1, V2, and C2 domains, rendering the assays difficult. For this reason, we studied antibody binding to only cathepsin L- and cathepsin D-treated molecules. The panel of MAbs used for this study included the 13H8 and 6E10 antibodies, which recognize the C4 and V2 domains, respectively, and the 1026 and 1088 antibodies, which recognize the V3 and V2 domains, respectively (66, 67). In addition, we included the V3 domain-specific MAb 447-52D (24), the broadly neutralizing, CD4-blocking MAb b12 (12, 70), and CD4-IgG (13). Of these, 6E10, b12, and CD4-IgG are known to bind to conformation-dependent sites involving sequences from several regions of the molecule.

Using a standard ELISA, we compared antibody binding to cathepsin L-treated and untreated MN-rgp120. The digestions ran to completion, as judged by the absence of intact gp120 when products were resolved by Coomassie blue-stained PAGE. We found that cathepsin L digestion of gp120 destroyed the ability to bind both the V3 domain-specific, virus-neutralizing 1026 and 447-52D MAbs and the C4 domain-specific, CD4-blocking 13H8 MAb. However, most of the binding of b12 and CD4-IgG was preserved, as well as much of the binding to the V2-specific antibodies 6E10 and 1088, although there appeared to be a significant reduction in binding affinity. This result can be explained by the fact that the two C4 sites, G_{431}-K_{432} and Y_{435}-A_{436}, and one V3 site, K_{327}-G_{328}, are located in close proximity to the epitopes recognized by the 13H8, 1026, and 447-52D MAbs, while the V2 domain is remote from the cathepsin L site. A different pattern of binding was observed with cathepsin D-treated gp120. In these experiments, binding to 13H8, 447-52D, and 1026 was preserved, although there appeared to be some reduction in binding affinity of the 1026 MAb. However, there was a large reduction in binding of the b12 and 6E10 MAbs and CD4-IgG. The inability of cathepsin D treatment to inhibit the binding of the 13H8 MAb can be attributed to the fact that the cathepsin D cleavage sites are located in the V2 and C2 domains, remote from the conformation-independent 13H8 epitope in the C4 domain. The large decrease in binding affinity of the b12 MAb and CD4-IgG for cathepsin D-treated gp120 might be explained by the fact that sequences in the C2 domain are known to be important for maintaining the structure of the CD4 binding site and that binding of the b12 MAb is dependent on contact sites in this region (55, 105). The sensitivity of 6E10 to cathepsin D inhibition compared to that of 1088 may reflect the fact that 6E10 is a conformation-dependent epitope that requires sequences in the C2, V3, and C3 domains, whereas 1088 is a conformation-independent epitope (G. R. Nakamura et al., unpublished results). Preservation of significant binding to 1088 in cathepsin D-treated gp120 suggested that despite cleavage between residues 274 and 275 in the C2 domain, the N-terminal and C-terminal portions of the envelope protein remain bound together in a noncovalent complex. Together, these results demonstrate that cathepsin cleavage sites are located in regions of gp120 recognized by neutralizing MAbs and CD4-IgG and that cleavage by cathepsins L and D differentially alters antibody and CD4 binding to these sites.

DISCUSSION

In these studies, we identified the locations of cleavage sites on MN-rgp120 recognized by three proteases (cathepsins L, S, and D) thought to be important in antigen processing and presentation. We found that these sites were not randomly distributed but, rather, occurred in regions of the envelope glycoprotein known to possess receptor binding sites and epitopes recognized by neutralizing antibodies. Comparative sequence analysis showed that the majority of these sites were highly conserved in the major clades of HIV, with some being conserved in both the chimpanzee form of HIV and SIV. Finally, we showed that cleavage by cathepsins L and D diminished the binding of neutralizing antibodies and CD4-IgG. We found that none of the experimentally determined cathepsin cleavage sites matched the cathepsin cleavage sites predicted by enzyme cleavage site prediction programs (see Tables S1 and S2 and Fig. S1 in the supplemental material), thus emphasizing the need for experimental studies. To some extent, the ability to predict cathepsin cleavage sites has been limited by the availability of experimental data, as evidenced by the paucity of cathepsin cleavage data in the MEROPS Peptidase Database (76). Moreover, there is uncertainty as to the extent to which cathepsin recognition sequences extend upstream and downstream of the scissile bond. The listing of N-terminal and C-terminal flanking sequences for the sites defined in this study (see Tables S1 and S2 in the supplemental material) will contribute to our knowledge of cathepsin recognition motifs.

Remarkably, seven of the eight cathepsin cleavage sites in gp120 identified in this study were located in regions of the envelope protein known to be associated with receptor binding or the binding of neutralizing monoclonal antibodies. For example, the V2 domain is known to contain epitopes recognized by virus neutralizing antibodies (36, 39, 62, 100) and has been termed the global regulator of virus neutralization (72). Moreover, the L_{181}-Y_{182} cathepsin D cleavage site is located just one amino acid away from the α4β7 receptor binding site (LDI/V) recently reported by Arthos et al. (2). The V3 domain is known as the principal neutralizing determinant, contains epitopes recognized by a variety of neutralizing antibodies (24, 66, 81), and is a key determinant of chemokine receptor tropism (14, 21, 101). Sequences in the C2 domain have been reported to be important for CD4 binding (55, 105), chemokine receptor binding (78, 79), and the binding of the broadly neutralizing b12 MAb (105). It is remarkable that one of the cathepsin S sites identified in the C2 domain is located adjacent to a chemokine receptor contact residue and that another is located at b12 MAb contact residues. The C4 domain is known to possess multiple contact residues for CD4 binding (55, 105), chemokine receptor binding (78, 79), and the binding of CD4-blocking neutralizing antibodies (56, 66, 67, 70). The importance of the CD4 binding site in antigen processing was noted previously (18, 95), where antibodies to the CD4 binding site inhibited cleavage by antigen processing enzymes and subsequent MHC class II antigen presentation. It is difficult to understand how this remarkable correspondence between receptor and neutralizing antibody binding sites and cathepsin cleavage sites could occur by chance. This is particularly significant in view of the fact that there are several regions in gp120 (C1, V1, C3, V4, and V5 domains) that appear to be devoid of cathepsin cleav-

age sites, indicating that cleavage sites are not randomly distributed.

The functional importance of the cathepsin cleavage sites identified above was further supported by the observation that six of the eight cathepsin cleavage sites in gp120 were highly conserved, with one (G_{431}-K_{432} in the C4 domain) found in HIV, HIV$_{cpz}$, and SIV and another (S_{261}-T_{262}) conserved in HIV and SIV. Previous studies have suggested that many viruses, including HIV, have evolved mechanisms to alter antigen processing to their advantage, as a way to escape or direct the immune response (103). Most of these mechanisms affect MHC class I-restricted cellular immune responses; however, mechanisms that alter MHC class II antigen presentation have also been reported (54).

HIV has developed a variety of mechanisms to evade the immune response. The hallmark of HIV infections is destruction of CD4 helper T cells, required to initiate and sustain effective antiviral immune responses (40) and to control virus replication (73, 80). Another mechanism to evade the immune response is the high level of sequence variation seen in all HIV proteins, particularly in the envelope protein, which typically exhibits multiple insertions and deletions. The virus also appears to have evolved epitope concealment mechanisms, especially in the unbound trimeric form of the envelope protein (prior to receptor ligation), that restrict access to antibody binding at neutralizing sites in the V3 domain, CD4 binding site, and membrane-proximal external region (MPER) (7, 34, 84, 90). Finally, the many N-linked glycosylation sites on gp120 are thought to form a protective "glycan shield" that provides yet another level of protection from the binding of neutralizing antibodies (98).

Our results suggest that HIV may have evolved another mechanism of immune escape involving insertion of protease cleavage sites in regions important for receptor binding and the binding of neutralizing antibodies. Cleavage at these sites may direct or modulate the immune response in a way that benefits survival of the virus. Because of the extraordinarily high level of sequence variation in HIV-1, resulting from high mutation and replication rates as well as immune selection, it is unlikely that these cleavage sites could be preserved unless they provided a significant fitness advantage for the virus. Recent studies by Tenzer et al. (92) suggested that the immunodominance of cytotoxic T-lymphocyte (CTL) epitopes is determined by proteasome digestion profiles and trimming by endoplasmic reticulum aminopeptidases. They further showed that CTL escape mutations involved amino acid substitutions that affected proteasome cleavages directly or sequences flanking cleavage sites in p17 and p24. The results from the present studies are consistent with the possibility that HIV might similarly regulate the immunodominance of MHC class II-restricted immune responses by tightly controlling proteolysis by antigen processing enzymes. The observation that the antigen processing sites are highly conserved is itself remarkable and consistent with this hypothesis. The additional observation that these sites are located in regions associated with receptor binding and neutralizing antibody binding is noteworthy and suggestive of important functional significance.

One potential explanation for the conservation of cathepsin cleavage sites at receptor binding sites is the fact that the receptor binding sites are among the few sites on virion-associated envelope proteins that are not protected by the glycan shield, and thus, they may be the only sites on the trimeric envelope protein accessible to proteases. However, it is unlikely that this can explain the data, since gp120 is readily shed from viruses and monomeric gp120 has multiple exposed regions that are not glycosylated. An alternative explanation may relate to an immune escape mechanism first described for poliovirus. Studies with poliovirus type 3 have shown that a major neutralizing epitope (antigenic site 1) contains a protease cleavage site and that cleavage at this site prevents the binding of neutralizing antibodies to an immunodominant epitope (63). It was suggested that this protease site may have evolved as a means by which the virus could escape from neutralizing antibodies directed to this site. The incorporation of protease cleavage sites at neutralizing epitopes, in effect, causes neutralizing epitopes to "self-destruct" after coming into contact with serum or cellular proteases. The effect of such cleavage could be to prevent either formation of neutralizing antibodies or binding of existing neutralizing antibodies to these sites. The possibility that key epitopes on the HIV envelope protein are labile and subject to destruction by extracellular proteases before they can stimulate antibody responses is intriguing. It could explain why it has been so difficult to elicit neutralizing antibodies with recombinant envelope proteins, despite the fact that they clearly possess the capacity to absorb bNAbs from HIV-positive sera (4, 59, 91). For this mechanism to be operative, two conditions must be met: cathepsin proteolysis must destroy the epitopes recognized by neutralizing antibodies, and the proteolysis must occur extracellularly, prior to binding of gp120 to antigen receptors on B cells. The antibody binding studies described in this paper showed that the binding to neutralizing antibodies and CD4-IgG was significantly reduced or eliminated by cathepsin cleavage, thus fulfilling the first condition. Since cathepsins are best known as lysosomal and endosomal enzymes, it was uncertain whether the second condition could be met. However, examination of the literature revealed that several cathepsins (e.g., cathepsins L, B, S, and K) can be secreted; these are known to play an important role in cancer biology, tissue remodeling, and inflammatory diseases (17, 57, 69, 102). The release of these enzymes has not been studied in the course of HIV infection or vaccination; however, cathepsin S has been reported to be secreted from activated macrophages (77) and is induced by treatment with adjuvants such as MF59 (65). Because cathepsin S is secreted and active at neutral pH and because cathepsin S sites are located at a variety of neutralizing epitopes in the C2, V3, and C4 domains, this enzyme appears to be a good candidate to mediate the destruction of neutralizing epitopes in gp120. While proteolysis of virion-associated envelope proteins would be expected to reduce virus infectivity, it is doubtful that this cleavage would be 100% effective. The high levels of plasma viremia and integrated provirus that occur in HIV infection would likely ensure that infection is sustained even if a large percentage of virus is inactivated by protease cleavage. Studies to assess the sensitivity of virions to cathepsin digestion are in progress. Several recent papers have suggested that HIV may enter cells by endocytosis as well as by virus-to-cell and cell-to-cell fusion (20, 49, 64). Although the relative importance of this pathway remains to be clarified, cleavage at the conserved cathepsin cleavage sites would be expected to occur

upon virion uptake; however, the implications for modulation of the immune response by a virus-infected cell are uncertain.

While the role of cathepsins in MHC class II immune responses is undisputed, they may also play an important role in MHC class I responses to HIV. A variety of MHC class I-restricted CTL epitopes and MHC class II-restricted T-cell epitopes occur at "hot spots" that appear to be located in close proximity (see Fig. S2 in the supplemental material) to the cathepsin cleavage sites identified in this paper (9, 33, 104). These include the cathepsin S sites in the C2 domain (35), the cathepsin S and L sites in the V3 domain (1, 82), and the cathepsin L sites in the C4 domain (42, 93). The colocation of these CTL epitopes with the cathepsin cleavage sites identified in this paper may result from the TAP-independent "cross-presentation" pathway that has been documented for dendritic cells and macrophages and is known to require cathepsin S (88). This pathway enables proteasome-independent MHC class I-restricted presentation of peptides generated by cathepsin S cleavage.

Identification of antigen processing sites promises to provide a new understanding of the molecular basis of the immune response to HIV envelope glycoprotein. Inactivation or relocation of cleavage sites recognized by antigen-processing proteases may provide a new approach to refocus the specificity of humoral and cellular responses to viral proteins. Although the functional target of bNAbs is the membrane-bound trimeric envelope protein on the surfaces of virions, we do not know whether the bNAbs in HIV-positive sera are elicited against virion-associated trimers or against monomeric envelope proteins shed into humoral circulation (37, 38, 68). Since a large percentage of bNAbs in HIV-positive sera can be adsorbed with monomeric gp120 (4, 59, 91) and since neither disulfide bond-stabilized trimers nor monomeric gp120 is effective in eliciting bNAbs (3, 51, 60), it appears that factors other than the oligomerization state must limit the neutralizing antibody response to HIV. Proteolytic cleavage of HIV envelope proteins may provide an explanation for why neither of these forms of the envelope protein has been effective in eliciting bNAbs.

Proteases are estimated to represent ~2% of the genes in the human genome (75), and it would not be surprising that HIV has evolved strategies to use cellular proteases to its advantage. The studies described here will contribute to our understanding of the specificity of antiviral immune responses and will add to our knowledge of the role of proteases in HIV biology.

REFERENCES

1. **Alexander-Miller, M. A., K. C. Parker, T. Tsukui, C. D. Pendleton, J. E. Coligan, and J. A. Berzofsky.** 1996. Molecular analysis of presentation by HLA-A2.1 of a promiscuously binding V3 loop peptide from the HIV-envelope protein to human cytotoxic T lymphocytes. Int. Immunol. **8:**641–649.
2. **Arthos, J., C. Cicala, E. Martinelli, K. Macleod, D. Van Ryk, D. Wei, Z. Xiao, T. D. Veenstra, T. P. Conrad, R. A. Lempicki, S. McLaughlin, M. Pascuccio, R. Gopaul, J. McNally, C. C. Cruz, N. Censoplano, E. Chung, K. N. Reitano, S. Kottilil, D. J. Goode, and A. S. Fauci.** 2008. HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. Nat. Immunol. **9:**301–309.
3. **Beddows, S., M. Franti, A. K. Dey, M. Kirschner, S. P. Iyer, D. C. Fisch, T. Ketas, E. Yuste, R. C. Desrosiers, P. J. Klasse, P. J. Maddon, W. C. Olson, and J. P. Moore.** 2007. A comparative immunogenicity study in rabbits of disulfide-stabilized, proteolytically cleaved, soluble trimeric human immunodeficiency virus type 1 gp140, trimeric cleavage-defective gp140 and monomeric gp120. Virology **360:**329–340.
4. **Binley, J. M., E. A. Lybarger, E. T. Crooks, M. S. Seaman, E. Gray, K. L. Davis, J. M. Decker, D. Wycuff, L. Harris, N. Hawkins, B. Wood, C. Nathe, D. Richman, G. D. Tomaras, F. Bibollet-Ruche, J. E. Robinson, L. Morris, G. M. Shaw, D. C. Montefiori, and J. R. Mascola.** 2008. Profiling the specificity of neutralizing antibodies in a large panel of plasmas from patients chronically infected with human immunodeficiency virus type 1 subtypes B and C. J. Virol. **82:**11651–11668.
5. **Binley, J. M., R. W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D. J. Anselma, P. J. Maddon, W. C. Olson, and J. P. Moore.** 2000. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. J. Virol. **74:**627–643.
6. **Blum, J. S., and P. Cresswell.** 1988. Role for intracellular proteases in the processing and transport of class II HLA antigens. Proc. Natl. Acad. Sci. USA **85:**3975–3979.
7. **Bou-Habib, D. C., G. Roderiquez, T. Oravecz, P. W. Berman, P. Lusso, and M. A. Norcross.** 1994. Cryptic nature of envelope V3 region epitopes protects primary monocytotropic human immunodeficiency virus type 1 from antibody neutralization. J. Virol. **68:**6006–6013.
8. **Boyd, S. E., R. N. Pike, G. B. Rudy, J. C. Whisstock, and M. Garcia de la Banda.** 2005. PoPS: a computational tool for modeling and predicting protease specificity. J. Bioinform. Comput. Biol. **3:**551–585.
9. **Brown, S. A., T. D. Lockey, C. Slaughter, K. S. Slobod, S. Surman, A. Zirkel, A. Mishra, V. R. Pagala, C. Coleclough, P. C. Doherty, and J. L. Hurwitz.** 2005. T cell epitope "hotspots" on the HIV type 1 gp120 envelope protein overlap with tryptic fragments displayed by mass spectrometry. AIDS Res. Hum. Retrovir. **21:**165–170.
10. **Buchbinder, S. P., D. V. Mehrotra, A. Duerr, D. W. Fitzgerald, R. Mogg, D. Li, P. B. Gilbert, J. R. Lama, M. Marmor, C. Del Rio, M. J. McElrath, D. R. Casimiro, K. M. Gottesdiener, J. A. Chodakewitz, L. Corey, and M. N. Robertson.** 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet **372:**1881–1893.
11. **Burton, D. R., R. C. Desrosiers, R. W. Doms, W. C. Koff, P. D. Kwong, J. P. Moore, G. J. Nabel, J. Sodroski, I. A. Wilson, and R. T. Wyatt.** 2004. HIV vaccine design and the neutralizing antibody problem. Nat. Immunol. **5:**233–236.
12. **Burton, D. R., J. Pyati, R. Koduri, S. J. Sharp, G. B. Thornton, P. W. Parren, L. S. Sawyer, R. M. Hendry, N. Dunlop, P. L. Nara, M. Lamacchia, E. Garratty, E. R. Stiehm, Y. J. Bryson, Y. Cao, J. Moore, D. D. Ho, and C. F. Barbas.** 1994. Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. Science **266:**1024–1027.
13. **Capon, D. J., S. M. Chamow, J. Mordenti, S. A. Marsters, T. Gregory, H. Mitsuya, R. A. Byrn, C. Lucas, F. M. Wurm, J. E. Groopman, S. Broder, and D. H. Smith.** 1989. Designing CD4 immunoadhesins for AIDS therapy. Nature **337:**525–531.
14. **Cardozo, T., T. Kimura, S. Philpott, B. Weiser, H. Burger, and S. Zolla-Pazner.** 2007. Structural basis for coreceptor selectivity by the HIV type 1 V3 loop. AIDS Res. Hum. Retrovir. **23:**415–426.
15. **Catanzaro, A. T., R. A. Koup, M. Roederer, R. T. Bailer, M. E. Enama, Z. Moodie, L. Gu, J. E. Martin, L. Novik, B. K. Chakrabarti, B. T. Butman, J. G. Gall, C. R. King, C. A. Andrews, R. Sheets, P. L. Gomez, J. R. Mascola, G. J. Nabel, and B. S. Graham.** 2006. Phase 1 safety and immunogenicity evaluation of a multiclade HIV-1 candidate vaccine delivered by a replication-defective recombinant adenovirus vector. J. Infect. Dis. **194:**1638–1649.
16. **Chapman, H. A.** 2006. Endosomal proteases in antigen presentation. Curr. Opin. Immunol. **18:**78–84.
17. **Chauhan, S. S., L. J. Goldstein, and M. M. Gottesman.** 1991. Expression of cathepsin L in human tumors. Cancer Res. **51:**1478–1481.
18. **Chien, P. C., Jr., S. Cohen, M. Tuen, J. Arthos, P. D. Chen, S. Patel, and C. E. Hioe.** 2004. Human immunodeficiency virus type 1 evades T-helper responses by exploiting antibodies that suppress antigen processing. J. Virol. **78:**7645–7652.
19. **Choe, Y., F. Leonetti, D. C. Greenbaum, F. Lecaille, M. Bogyo, D. Bromme, J. A. Ellman, and C. S. Craik.** 2006. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. J. Biol. Chem. **281:**12824–12832.
20. **Clotet-Codina, I., B. Bosch, J. Senserrich, M. T. Fernandez-Figueras, R. Pena, E. Ballana, M. Bofill, B. Clotet, and J. A. Este.** 2009. HIV endocytosis after dendritic cell to T cell viral transfer leads to productive virus infection. Antiviral Res. **83:**94–98.
21. **Cocchi, F., A. L. DeVico, A. Garzino-Demo, A. Cara, R. C. Gallo, and P.**

Lusso. 1996. The V3 domain of the HIV-1 gp120 envelope glycoprotein is critical for chemokine-mediated blockade of infection. Nat. Med. **2:**1244–1247.

22. **Cohen, J.** 2007. AIDS research. Did Merck's failed HIV vaccine cause harm? Science **318:**1048–1049.

23. **Cohen, J.** 2007. AIDS research. Promising AIDS vaccine's failure leaves field reeling. Science **318:**28–29.

24. **Conley, A. J., M. K. Gorny, J. A. Kessler II, L. J. Boots, M. Ossorio-Castro, S. Koenig, D. W. Lineberger, E. A. Emini, C. Williams, and S. Zolla-Pazner.** 1994. Neutralization of primary human immunodeficiency virus type 1 isolates by the broadly reactive anti-V3 monoclonal antibody, 447-52D. J. Virol. **68:**6994–7000.

25. **Decker, J. M., F. Bibollet-Ruche, X. Wei, S. Wang, D. N. Levy, W. Wang, E. Delaporte, M. Peeters, C. A. Derdeyn, S. Allen, E. Hunter, M. S. Saag, J. A. Hoxie, B. H. Hahn, P. D. Kwong, J. E. Robinson, and G. M. Shaw.** 2005. Antigenic conservation and immunogenicity of the HIV coreceptor binding site. J. Exp. Med. **201:**1407–1419.

26. **Dey, B., K. Svehla, L. Xu, D. Wycuff, T. Zhou, G. Voss, A. Phogat, B. K. Chakrabarti, Y. Li, G. Shaw, P. D. Kwong, G. J. Nabel, J. R. Mascola, and R. T. Wyatt.** 2009. Structure-based stabilization of HIV-1 gp120 enhances humoral immune responses to the induced co-receptor binding site. PLoS Pathog. **5:**e1000445.

27. **Douek, D. C., P. D. Kwong, and G. J. Nabel.** 2006. The rational design of an AIDS vaccine. Cell **124:**677–681.

28. **Driessen, C., R. A. Bryant, A. M. Lennon-Dumenil, J. A. Villadangos, P. W. Bryant, G. P. Shi, H. A. Chapman, and H. L. Ploegh.** 1999. Cathepsin S controls the trafficking and maturation of MHC class II molecules in dendritic cells. J. Cell Biol. **147:**775–790.

29. **Egan, M. A., S. Y. Chong, S. Megati, D. C. Montefiori, N. F. Rose, J. D. Boyer, M. K. Sidhu, J. Quiroz, M. Rosati, E. B. Schadeck, G. N. Pavlakis, D. B. Weiner, J. K. Rose, Z. R. Israel, S. A. Udem, and J. H. Eldridge.** 2005. Priming with plasmid DNAs expressing interleukin-12 and simian immunodeficiency virus gag enhances the immunogenicity and efficacy of an experimental AIDS vaccine based on recombinant vesicular stomatitis virus. AIDS Res. Hum. Retrovir. **21:**629–643.

30. **Ellenberger, D., R. A. Otten, B. Li, M. Aidoo, I. V. Rodriguez, C. A. Sariol, M. Martinez, M. Monsour, L. Wyatt, M. G. Hudgens, E. Kraisilburd, B. Moss, H. Robinson, T. Folks, and S. Butera.** 2006. HIV-1 DNA/MVA vaccination reduces the per exposure probability of infection during repeated mucosal SHIV challenges. Virology **352:**216–225.

31. **Fauci, A. S., M. I. Johnston, C. W. Dieffenbach, D. R. Burton, S. M. Hammer, J. A. Hoxie, M. Martin, J. Overbaugh, D. I. Watkins, A. Mahmoud, and W. C. Greene.** 2008. HIV vaccine research: the way forward. Science **321:**530–532.

32. **Flynn, N. M., D. N. Forthal, C. D. Harro, F. N. Judson, K. H. Mayer, and M. F. Para.** 2005. Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. J. Infect. Dis. **191:**654–665.

33. **Frahm, N., B. Baker, and C. Brander.** 2008. Identification and optimal definition of HIV-derived cytotoxic lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution, p. 3–24. *In* B. Korber, C. Brander, B. Haynes, R. Koup, J. Moore, B. Walker, and B. A. Watkins (ed.), HIV molecular immunology 2008. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM.

34. **Frey, G., H. Peng, S. Rits-Volloch, M. Morelli, Y. Cheng, and B. Chen.** 2008. A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. Proc. Natl. Acad. Sci. USA **105:**3739–3744.

35. **Fukada, K., H. Tomiyama, Y. Chujoh, K. Miwa, Y. Kaneko, S. Oka, and M. Takiguchi.** 1999. HLA-A*1101-restricted cytotoxic T lymphocyte recognition for a novel epitope derived from the HIV-1 Env protein. AIDS **13:**2597–2599.

36. **Fung, M. S., C. R. Sun, W. L. Gordon, R. S. Liou, T. W. Chang, W. N. Sun, E. S. Daar, and D. D. Ho.** 1992. Identification and characterization of a neutralization site within the second variable region of human immunodeficiency virus type 1 gp120. J. Virol. **66:**848–856.

37. **Gilbert, M., J. Kirihara, and J. Mills.** 1991. Enzyme-linked immunoassay for human immunodeficiency virus type 1 envelope glycoprotein 120. J. Clin. Microbiol. **29:**142–147.

38. **Gilbert, P. B., Y. L. Chiu, M. Allen, D. N. Lawrence, C. Chapdu, H. Israel, D. Holman, M. C. Keefer, M. Wolff, and S. E. Frey.** 2003. Long-term safety analysis of preventive HIV-1 vaccines evaluated in AIDS vaccine evaluation group NIAID-sponsored phase I and II clinical trials. Vaccine **21:**2933–2947.

39. **Gorny, M. K., L. Stamatatos, B. Volsky, K. Revesz, C. Williams, X. H. Wang, S. Cohen, R. Staudinger, and S. Zolla-Pazner.** 2005. Identification of a new quaternary neutralizing epitope on human immunodeficiency virus type 1 virus particles. J. Virol. **79:**5232–5237.

40. **Grundner, C., M. Pancera, J. M. Kang, M. Koch, J. Sodroski, and R. Wyatt.** 2004. Factors limiting the immunogenicity of HIV-1 gp120 envelope glycoproteins. Virology **330:**233–248.

41. **Gurgo, C., H. G. Guo, G. Franchini, A. Aldovini, E. Collalti, K. Farrell, F. Wong-Staal, R. C. Gallo, and M. S. Reitz, Jr.** 1988. Envelope sequences of two new United States HIV-1 isolates. Virology **164:**531–536.

42. **Harrer, T., E. Harrer, S. A. Kalams, P. Barbosa, A. Trocha, R. P. Johnson, T. Elbeik, M. B. Feinberg, S. P. Buchbinder, and B. D. Walker.** 1996. Cytotoxic T lymphocytes in asymptomatic long-term nonprogressing HIV-1 infection. Breadth and specificity of the response and relation to in vivo viral quasispecies in a person with prolonged infection and low viral load. J. Immunol. **156:**2616–2623.

43. **Haynes, B., and D. Montefiori.** 2006. Aiming to induce broadly reactive neutralizing antibody responses with HIV-1 vaccine candidates. Expert Rev. Vaccines **5:**579–595.

44. **Hsieh, C. S., P. deRoos, K. Honey, C. Beers, and A. Y. Rudensky.** 2002. A role for cathepsin L and cathepsin S in peptide generation for MHC class II presentation. J. Immunol. **168:**2618–2625.

45. **Hsing, L. C., and A. Y. Rudensky.** 2005. The lysosomal cysteine proteases in MHC class II antigen presentation. Immunol. Rev. **207:**229–241.

46. **Hu, S. L., and L. Stamatatos.** 2007. Prospects of HIV Env modification as an approach to HIV vaccine design. Curr. HIV Res. **5:**507–513.

47. **Huang, C. C., F. Stricher, L. Martin, J. M. Decker, S. Majeed, P. Barthe, W. A. Hendrickson, J. Robinson, C. Roumestand, J. Sodroski, R. Wyatt, G. M. Shaw, C. Vita, and P. D. Kwong.** 2005. Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. Structure **13:**755–768.

48. **Huang, C. C., M. Tang, M. Y. Zhang, S. Majeed, E. Montabana, R. L. Stanfield, D. S. Dimitrov, B. Korber, J. Sodroski, I. A. Wilson, R. Wyatt, and P. D. Kwong.** 2005. Structure of a V3-containing HIV-1 gp120 core. Science **310:**1025–1028.

49. **Hubner, W., G. P. McNerney, P. Chen, B. M. Dale, R. E. Gordon, F. Y. Chuang, X. D. Li, D. M. Asmuth, T. Huser, and B. K. Chen.** 2009. Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science **323:**1743–1747.

50. **Johnston, M. I., and A. S. Fauci.** 2007. An HIV vaccine—evolving concepts. N. Engl. J. Med. **356:**2073–2081.

51. **Kang, Y. K., S. Andjelic, J. M. Binley, E. T. Crooks, M. Franti, S. P. Iyer, G. P. Donovan, A. K. Dey, P. Zhu, K. H. Roux, R. J. Durso, T. F. Parsons, P. J. Maddon, J. P. Moore, and W. C. Olson.** 2009. Structural and immunogenicity studies of a cleaved, stabilized envelope trimer derived from subtype A HIV-1. Vaccine **27:**5120–5132.

52. **Katoh, K., K. Kuma, H. Toh, and T. Miyata.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. **33:**511–518.

53. **Keele, B. F., E. E. Giorgi, J. F. Salazar-Gonzalez, J. M. Decker, K. T. Pham, M. G. Salazar, C. Sun, T. Grayson, S. Wang, H. Li, X. Wei, C. Jiang, J. L. Kirchherr, F. Gao, J. A. Anderson, L. H. Ping, R. Swanstrom, G. D. Tomaras, W. A. Blattner, P. A. Goepfert, J. M. Kilby, M. S. Saag, E. L. Delwart, M. P. Busch, M. S. Cohen, D. C. Montefiori, B. F. Haynes, B. Gaschen, G. S. Athreya, H. Y. Lee, N. Wood, C. Seoighe, A. S. Perelson, T. Bhattacharya, B. T. Korber, B. H. Hahn, and G. M. Shaw.** 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. USA **105:**7552–7557.

54. **Kessler, T., M. Reich, G. Jahn, E. Tolosa, A. Beck, H. Kalbacher, H. Overkleeft, S. Schempp, and C. Driessen.** 2008. Human cytomegalovirus infection interferes with major histocompatibility complex type II maturation and endocytic proteases in dendritic cells at multiple levels. J. Gen. Virol. **89:**2427–2436.

55. **Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson.** 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature **393:**648–659.

56. **Lasky, L. A., G. Nakamura, D. H. Smith, C. Fennie, C. Shimasaki, E. Patzer, P. Berman, T. Gregory, and D. J. Capon.** 1987. Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. Cell **50:**975–985.

57. **Lemaire, R., G. Huet, F. Zerimech, G. Grard, C. Fontaine, B. Duquesnoy, and R. M. Flipo.** 1997. Selective induction of the secretion of cathepsins B and L by cytokines in synovial fibroblast-like cells. Br. J. Rheumatol. **36:**735–743.

58. **Leonard, C. K., M. W. Spellman, L. Riddle, R. J. Harris, J. N. Thomas, and T. J. Gregory.** 1990. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J. Biol. Chem. **265:**10373–10382.

59. **Li, Y., S. A. Migueles, B. Welcher, K. Svehla, A. Phogat, M. K. Louder, X. Wu, G. M. Shaw, M. Connors, R. T. Wyatt, and J. R. Mascola.** 2007. Broad HIV-1 neutralization mediated by CD4-binding site antibodies. Nat. Med. **13:**1032–1034.

60. **Li, Y., K. Svehla, N. L. Mathy, G. Voss, J. R. Mascola, and R. Wyatt.** 2006. Characterization of antibody responses elicited by human immunodeficiency virus type 1 primary isolate trimeric and monomeric envelope glycoproteins in selected adjuvants. J. Virol. **80:**1414–1426.

61. **McElrath, M. J., S. C. De Rosa, Z. Moodie, S. Dubey, L. Kierstead, H. Janes, O. D. Defawe, D. K. Carter, J. Hural, R. Akondy, S. P. Buch-

binder, M. N. Robertson, D. V. Mehrotra, S. G. Self, L. Corey, J. W. Shiver, and D. R. Casimiro. 2008. HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case-cohort analysis. Lancet 372:1894–1905.

62. McKeating, J. A., C. Shotton, J. Cordell, S. Graham, P. Balfe, N. Sullivan, M. Charles, M. Page, A. Bolmstedt, S. Olofsson, et al. 1993. Characterization of neutralizing monoclonal antibodies to linear and conformation-dependent epitopes within the first and second variable domains of human immunodeficiency virus type 1 gp120. J. Virol. 67:4932–4944.

63. Minor, P. D., M. Ferguson, A. Phillips, D. I. Magrath, A. Huovilainen, and T. Hovi. 1987. Conservation in vivo of protease cleavage sites in antigenic sites of poliovirus. J. Gen. Virol. 68:1857–1865.

64. Miyauchi, K., Y. Kim, O. Latinovic, V. Morozov, and G. B. Melikyan. 2009. HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. Cell 137:433–444.

65. Mosca, F., E. Tritto, A. Muzzi, E. Monaci, F. Bagnoli, C. Iavarone, D. O'Hagan, R. Rappuoli, and E. De Gregorio. 2008. Molecular and cellular signatures of human vaccine adjuvants. Proc. Natl. Acad. Sci. USA 105:10501–10506.

66. Nakamura, G. R., R. Byrn, K. Rosenthal, J. P. Porter, M. R. Hobbs, L. Riddle, D. J. Eastman, D. Dowbenko, T. Gregory, B. M. Fendly, et al. 1992. Monoclonal antibodies to the extracellular domain of HIV-1IIIB gp160 that neutralize infectivity, block binding to CD4, and react with diverse isolates. AIDS Res. Hum. Retrovir. 8:1875–1885.

67. Nakamura, G. R., R. Byrn, D. M. Wilkes, J. A. Fox, M. R. Hobbs, R. Hastings, H. C. Wessling, M. A. Norcross, B. M. Fendly, and P. W. Berman. 1993. Strain specificity and binding affinity requirements of neutralizing monoclonal antibodies to the C4 domain of gp120 from human immunodeficiency virus type 1. J. Virol. 67:6179–6191.

68. Oh, S. K., W. W. Cruikshank, J. Raina, G. C. Blanchard, W. H. Adler, J. Walker, and H. Kornfeld. 1992. Identification of HIV-1 envelope glycoprotein in the serum of AIDS and ARC patients. J. Acquir. Immune Defic. Syndr. 5:251–256.

69. Palermo, C., and J. A. Joyce. 2008. Cysteine cathepsin proteases as pharmacological targets in cancer. Trends Pharmacol. Sci. 29:22–28.

70. Pantophlet, R., and D. R. Burton. 2006. GP120: target for neutralizing HIV-1 antibodies. Annu. Rev. Immunol. 24:739–769.

71. Phan, K. O., M. E. Callahan, S. Vanichseni, C. J. Du, S. Raktham, N. Young, K. Choopanya, T. D. Mastro, and S. Subbarao. 2000. A comparison of full-length glycoprotein 120 from incident HIV type 1 subtype E and B infections in Bangkok injecting drug users with prototype E and B strains that are components of a candidate vaccine. AIDS Res. Hum. Retrovir. 16:1445–1450.

72. Pinter, A., W. J. Honnen, Y. He, M. K. Gorny, S. Zolla-Pazner, and S. C. Kayman. 2004. The V1/V2 domain of gp120 is a global regulator of the sensitivity of primary human immunodeficiency virus type 1 isolates to neutralization by antibodies commonly induced upon infection. J. Virol. 78:5205–5215.

73. Pitcher, C. J., C. Quittner, D. M. Peterson, M. Connors, R. A. Koup, V. C. Maino, and L. J. Picker. 1999. HIV-1-specific CD4+ T cells are detectable in most individuals with active HIV-1 infection, but decline with prolonged viral suppression. Nat. Med. 5:518–525.

74. Pitisuttithum, P., P. Gilbert, M. Gurwith, W. Heyward, M. Martin, F. van Griensven, D. Hu, J. W. Tappero, and K. Choopanya. 2006. Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. J. Infect. Dis. 194:1661–1671.

75. Puente, X. S., L. M. Sanchez, C. M. Overall, and C. Lopez-Otin. 2003. Human and mouse proteases: a comparative genomic approach. Nat. Rev. Genet. 4:544–558.

76. Rawlings, N. D., F. R. Morton, C. Y. Kok, J. Kong, and A. J. Barrett. 2008. MEROPS: the peptidase database. Nucleic Acids Res. 36:D320–D325.

77. Reddy, V. Y., Q. Y. Zhang, and S. J. Weiss. 1995. Pericellular mobilization of the tissue-destructive cysteine proteinases, cathepsins B, L, and S, by human monocyte-derived macrophages. Proc. Natl. Acad. Sci. USA 92:3849–3853.

78. Rizzuto, C., and J. Sodroski. 2000. Fine definition of a conserved CCR5-binding region on the human immunodeficiency virus type 1 glycoprotein 120. AIDS Res. Hum. Retrovir. 16:741–749.

79. Rizzuto, C. D., R. Wyatt, N. Hernandez-Ramos, Y. Sun, P. D. Kwong, W. A. Hendrickson, and J. Sodroski. 1998. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. Science 280:1949–1953.

80. Rosenberg, E. S., J. M. Billingsley, A. M. Caliendo, S. L. Boswell, P. E. Sax, S. A. Kalams, and B. D. Walker. 1997. Vigorous HIV-1-specific CD4+ T cell responses associated with control of viremia. Science 278:1447–1450.

81. Rusche, J. R., K. Javaherian, C. McDanal, J. Petro, D. L. Lynn, R. Grimaila, A. Langlois, R. C. Gallo, L. O. Arthur, P. J. Fischinger, et al. 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind a 24-amino acid sequence of the viral envelope, gp120. Proc. Natl. Acad. Sci. USA 85:3198–3202.

82. Sabbaj, S., A. Bansal, G. D. Ritter, C. Perkins, B. H. Edwards, E. Gough, J. Tang, J. J. Szinger, B. Korber, C. M. Wilson, R. A. Kaslow, M. J. Mulligan, and P. A. Goepfert. 2003. Cross-reactive CD8+ T cell epitopes identified in US adolescent minorities. J. Acquir. Immune Defic. Syndr. 33:426–438.

83. Santra, S., M. S. Seaman, L. Xu, D. H. Barouch, C. I. Lord, M. A. Lifton, D. A. Gorgone, K. R. Beaudry, K. Svehla, B. Welcher, B. K. Chakrabarti, Y. Huang, Z. Y. Yang, J. R. Mascola, G. J. Nabel, and N. L. Letvin. 2005. Replication-defective adenovirus serotype 5 vectors elicit durable cellular and humoral immune responses in nonhuman primates. J. Virol. 79:6516–6522.

84. Saunders, C. J., R. A. McCaffrey, I. Zharkikh, Z. Kraft, S. E. Malenbaum, B. Burke, C. Cheng-Mayer, and L. Stamatatos. 2005. The V1, V2, and V3 regions of the human immunodeficiency virus type 1 envelope differentially affect the viral phenotype in an isolate-dependent manner. J. Virol. 79:9069–9080.

85. Scarborough, P. E., K. Guruprasad, C. Topham, G. R. Richo, G. E. Conner, T. L. Blundell, and B. M. Dunn. 1993. Exploration of subsite binding specificity of human cathepsin D through kinetics and rule-based molecular modeling. Protein Sci. 2:264–276.

86. Schechter, I., and A. Berger. 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27:157–162.

87. Schulke, N., M. S. Vesanen, R. W. Sanders, P. Zhu, M. Lu, D. J. Anselma, A. R. Villa, P. W. Parren, J. M. Binley, K. H. Roux, P. J. Maddon, J. P. Moore, and W. C. Olson. 2002. Oligomeric and conformational properties of a proteolytically mature, disulfide-stabilized human immunodeficiency virus type 1 gp140 envelope glycoprotein. J. Virol. 76:7760–7776.

88. Shen, L., L. J. Sigal, M. Boes, and K. L. Rock. 2004. Important role of cathepsin S in generating peptides for TAP-independent MHC class I crosspresentation in vivo. Immunity 21:155–165.

89. Srivastava, I. K., E. Kan, Y. Sun, V. A. Sharma, J. Cisto, B. Burke, Y. Lian, S. Hilt, Z. Biron, K. Hartog, L. Stamatatos, R. Diaz-Avalos, R. H. Cheng, J. B. Ulmer, and S. W. Barnett. 2008. Comparative evaluation of trimeric envelope glycoproteins derived from subtype C and B HIV-1 R5 isolates. Virology 372:273–290.

90. Stamatatos, L., and C. Cheng-Mayer. 1998. An envelope modification that renders a primary, neutralization-resistant clade B human immunodeficiency virus type 1 isolate highly susceptible to neutralization by sera from other clades. J. Virol. 72:7840–7845.

91. Steimer, K. S., C. J. Scandella, P. V. Skiles, and N. L. Haigwood. 1991. Neutralization of divergent HIV-1 isolates by conformation-dependent human antibodies to Gp120. Science 254:105–108.

92. Tenzer, S., E. Wee, A. Burgevin, G. Stewart-Jones, L. Friis, K. Lamberth, C. H. Chang, M. Harndahl, M. Weimershaus, J. Gerstoft, N. Akkad, P. Klenerman, L. Fugger, E. Y. Jones, A. J. McMichael, S. Buus, H. Schild, P. van Endert, and A. K. Iversen. 2009. Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat. Immunol. 10:636–646.

93. Tomiyama, H., T. Sakaguchi, K. Miwa, S. Oka, A. Iwamoto, Y. Kaneko, and M. Takiguchi. 1999. Identification of multiple HIV-1 CTL epitopes presented by HLA-B*5101 molecules. Hum. Immunol. 60:177–186.

94. Trombetta, E. S., and I. Mellman. 2005. Cell biology of antigen processing in vitro and in vivo. Annu. Rev. Immunol. 23:975–1028.

95. Tuen, M., M. L. Visciano, P. C. Chien, Jr., S. Cohen, P. D. Chen, J. Robinson, Y. He, A. Pinter, M. K. Gorny, and C. E. Hioe. 2005. Characterization of antibodies that inhibit HIV gp120 antigen processing and presentation. Eur. J. Immunol. 35:2541–2551.

96. van Noort, J. M., and A. C. van der Drift. 1989. The selectivity of cathepsin D suggests an involvement of the enzyme in the generation of T-cell epitopes. J. Biol. Chem. 264:14159–14164.

97. Villadangos, J. A., and H. L. Ploegh. 2000. Proteolysis in MHC class II antigen presentation: who's in charge? Immunity 12:233–239.

98. Wei, X., J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, and G. M. Shaw. 2003. Antibody neutralization and escape by HIV-1. Nature 422:307–312.

99. Wolf, P. R., and H. L. Ploegh. 1995. How MHC class II molecules acquire peptide cargo: biosynthesis and trafficking through the endocytic pathway. Annu. Rev. Cell Dev. Biol. 11:267–306.

100. Wu, Z., S. C. Kayman, W. Honnen, K. Revesz, H. Chen, S. Vijh-Warrier, S. A. Tilley, J. McKeating, C. Shotton, and A. Pinter. 1995. Characterization of neutralization epitopes in the V2 region of human immunodeficiency virus type 1 gp120: role of glycosylation in the correct folding of the V1/V2 domain. J. Virol. 69:2271–2278.

101. Xiao, L., S. M. Owen, I. Goldman, A. A. Lal, J. J. deJong, J. Goudsmit, and R. B. Lal. 1998. CCR5 coreceptor usage of non-syncytium-inducing primary HIV-1 is independent of phylogenetically distinct global HIV-1 isolates: delineation of consensus motif in the V3 domain that predicts CCR5 usage. Virology 240:83–92.

102. Yasuda, Y., J. Kaleta, and D. Bromme. 2005. The role of cathepsins in

osteoporosis and arthritis: rationale for the design of new therapeutics. Adv. Drug Deliv. Rev. **57:**973–993.

103. **Yewdell, J. W., and J. R. Bennink.** 1999. Mechanisms of viral interference with MHC class I antigen processing and presentation. Annu. Rev. Cell Dev. Biol. **15:**579–606.

104. **Zhan, X., J. L. Hurwitz, S. A. Brown, and K. S. Slobod.** 2007. HIV-1 envelope T cell epitope "hotspots" among mice and humans and among

CD4+ and CD8+ T cell subpopulations. AIDS Res. Hum. Retrovir. **23:** 471–476.

105. **Zhou, T., L. Xu, B. Dey, A. J. Hessell, D. Van Ryk, S. H. Xiang, X. Yang, M. Y. Zhang, M. B. Zwick, J. Arthos, D. R. Burton, D. S. Dimitrov, J. Sodroski, R. Wyatt, G. J. Nabel, and P. D. Kwong.** 2007. Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature **445:** 732–737.