

Structure of the Core of the HIV-1 gp120 Exterior Envelope Glycoprotein

Richard Wyatt,¹ Peter D. Kwong,² Wayne A. Hendrickson,^{2,3} and Joseph G. Sodroski^{1,4}

¹*Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, MA 02115*

²*Department of Pathology, Harvard Medical School, Boston, MA 02115*

³*Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032*

⁴*Howard Hughes Medical Institute, Columbia University, New York, NY 10032*

The x-ray crystal structure of the gp120 core derived from the HXBc2 strain of HIV-1, in a ternary complex with two-domain CD4 and a neutralizing antibody, has been solved [1]. The gp120 core contains deletions of the V1, V2 and V3 variable loops and of the N- and C-termini, compared with full-length gp120. The implications of this structure for understanding the interaction of the HIV-1 envelope glycoproteins with the viral receptors, CD4 and chemokine receptors, and with the humoral immune response of the infected host have been reported [1–3]. A general review of the structure and function of the HIV-1 envelope glycoproteins, taking into account the x-ray crystallographic data, has been published [4]. Here we present illustrations that allow the reader to locate individual gp120 residues on the three-dimensional gp120 core structure and that summarize our current understanding of the discontinuous gp120 regions involved in ligand interactions.

Figure 1 shows a stereo diagram of an alpha-carbon trace of the gp120 core, in its CD4-bound state. The view shown is from the perspective of the CD4 molecule. Every tenth C α is marked with a filled circle, and every twentieth residue is labeled. Disulfide connections are depicted in a ball-and-stick format. Shown are the ordered residues, 90–396 and 410–492.

Beneath the stereo diagram in Figure 1 is a structure-based sequence alignment. Shown are the sequences of “HIV-1 B” (core gp120 from clade B, strain HXBc2 used in the crystallographic studies), “C” (HIV-1 clade C, strain

UG268A2), “O” (HIV-1 clade O, strain ANT70), “HIV-2” (strain ROD), and “SIY” (African green monkey isolate, clone GR1-1). The secondary structure assignments are shown as arrows and cylinders, with (x) denoting residues which are disordered in the present structure. The ‘gags’ sequence at the N-terminus and the ‘gag’ sequence in the V1/V2 and V3 loops are conserved by an open circle truncation. Solvent accessibility is indicated for each residue by an open circle if the fractional solvent accessibility is greater than 0.4, a half-closed circle if 0.1 to 0.4, and a closed circle if less than 0.1. Sequence variability observed among primate immunodeficiency viruses is indicated below the solvent accessibility by the number of horizontal hash marks: 1 mark, residues conserved among all primate immunodeficiency viruses; 2 marks, conserved among all HIV-1 isolates; 3 marks, exhibits moderate variation among HIV-1 isolates; and 4 marks, exhibits significant variability among HIV-1 isolates. In assessing conservation, all single atom changes were permitted as well as larger substitutions if the character of the sidechain was conserved (e.g., K to R or F to L). N-linked glycosylation is indicated by “n” for the high mannose additions and “c” for the complex additions observed in mammalian cells [5]. Residues of gp120 in direct contact with CD4 are indicated by “*”. Direct contact is a more restrictive criterion of interaction than the often used loss of solvent accessible surface; residues of gp120 which show loss of solvent accessible surface but are not in direct contact are 123, 124, 126, 257, 278, 282, 364, 471, 475, 476 and 477.

Figure 2 shows a series of ribbon diagrams of the gp120 core, again from the perspective of the CD4 glycoprotein. In Figure 2a, the inner and outer domains are shown in red and yellow, respectively, and the bridging sheet is colored blue. The inner domain is believed to interact with the gp41 envelope glycoprotein, while the outer domain, which is quite variable and heavily glycosylated (Figure 2b), is believed to be exposed on the assembled envelope glycoprotein trimer [3]. The “proximal” side of the gp120 core, which includes the N- and C-termini, is believed to reside near the viral membrane. The “distal” side of gp120 is believed to face the target cell membrane after CD4 binding occurs. The N- and C-termini, the structures related to some of the variable loops, and some of the secondary structural elements are labelled.

IV-1

DEC 98

In Figure 2b, the asparagine residues (including side chains) modified by N-linked glycosylation in mammalian cells [5] are shown in blue. Note the concentration of glycosylation on the gp120 surface believed to face outward on the trimeric envelope glycoprotein spike.

Figure 2c shows the gp120 residues (in red) that are within 4 Angstroms of CD4 in the complex. Residues shown by mutagenesis to be important for binding of gp120-CD4 complexes to the chemokine receptor [2] are colored green. The V3 loop (not shown) also contributes to chemokine receptor binding.

Figure 2d shows the gp120 residues implicated in binding neutralizing antibodies, either those directed against the CD4 binding site (CD4BS) (red), those directed against the CD4-induced (CD4i) epitopes (green), or the 2G12 antibody (blue). The CD4BS antibodies block CD4 binding and are represented by the F105 and 15e antibodies. The CD4i antibodies recognize elements of the bridging sheet and block binding of gp120-CD4 complexes to chemokine receptors. The 2G12 antibody recognizes carbohydrate-dependent structures on the gp120 outer domain. Few known antibodies recognize this heavily glycosylated gp120 outer domain.

References

- 1 Kwong PD, Wyatt R, Robinson J, Sweet R, Sodroski J and Hendrickson W. Structure of an HIV-1 gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, (1998) **393**:649–59.
- 2 Rizzuto C, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA and Sodroski J. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science*, (1998) **280**:1949–53.
- 3 Wyatt R, Kwong PD, Desjardins E, Sweet R, Robinson J, Hendrickson W and Sodroski J. The antigenic structure of the human immunodeficiency virus gp120 envelope glycoprotein. *Nature*, (1998) **393**:705–11.
- 4 Wyatt R and Sodroski J. The human immunodeficiency virus (HIV-1) envelope glycoproteins: Fusogens, antigens and immunogens. *Science*, (1998) **280**:1884–8.
- 5 Leonard C, Spellman M, Riddle L, Harris R, Thomas J and Gregory T. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J. Biol. Chem.*, (1990) **265**:10373–82.

Figure 1. Stereo diagram of an alpha-carbon trace of the gp120 core. Reprinted from, Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. 1998 Jun 18;393(6686):648–59. by permission from *Nature* copyright Macmillan Magazines Ltd.

IV-3

DEC 98

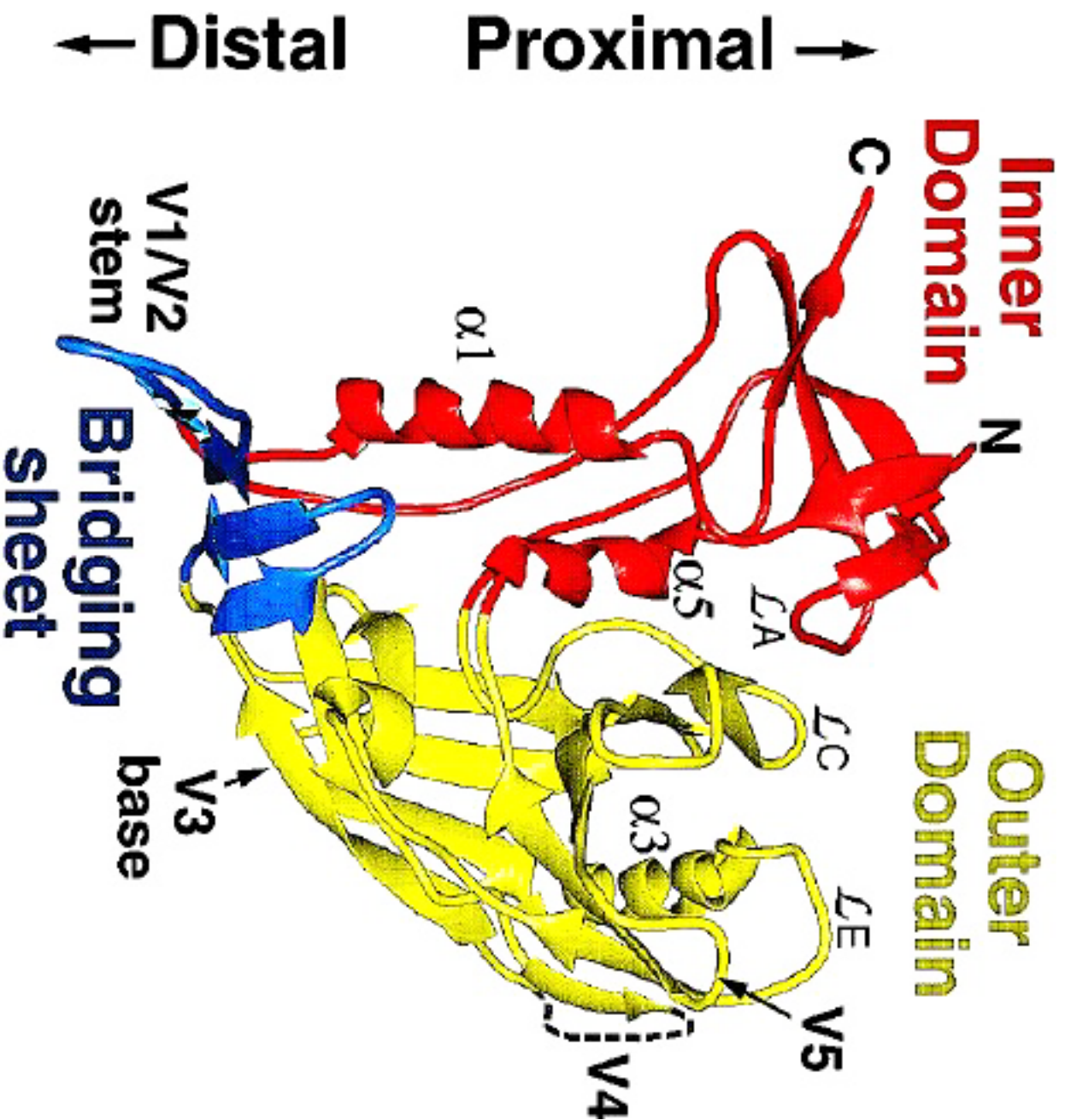


Figure 2a. gp120 core; inner domain (red), outer domain (yellow), bridging sheet (blue).

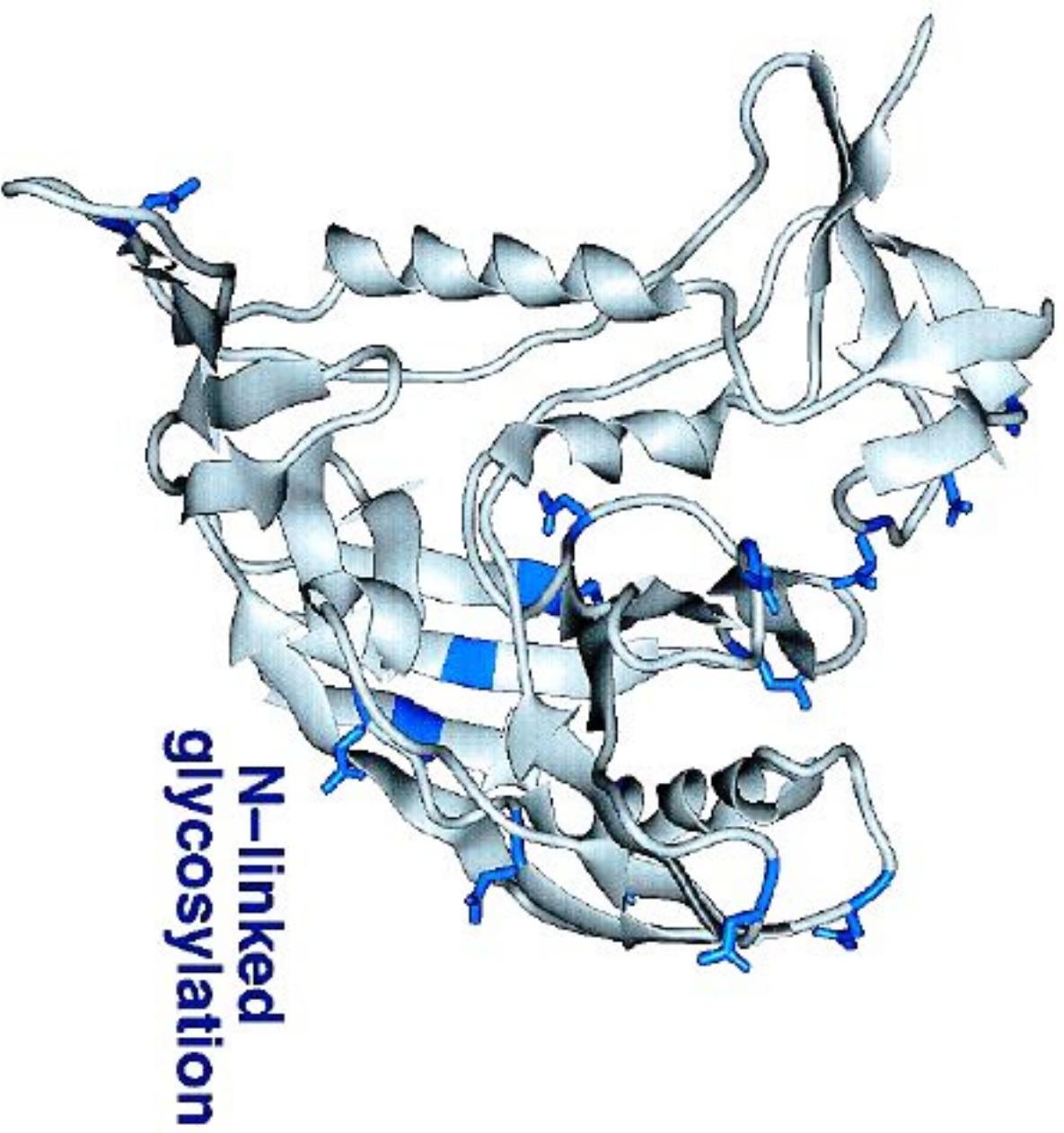


Figure 2b. gp120 core; asparagine residues modified by N-linked glycosylation shown in blue.

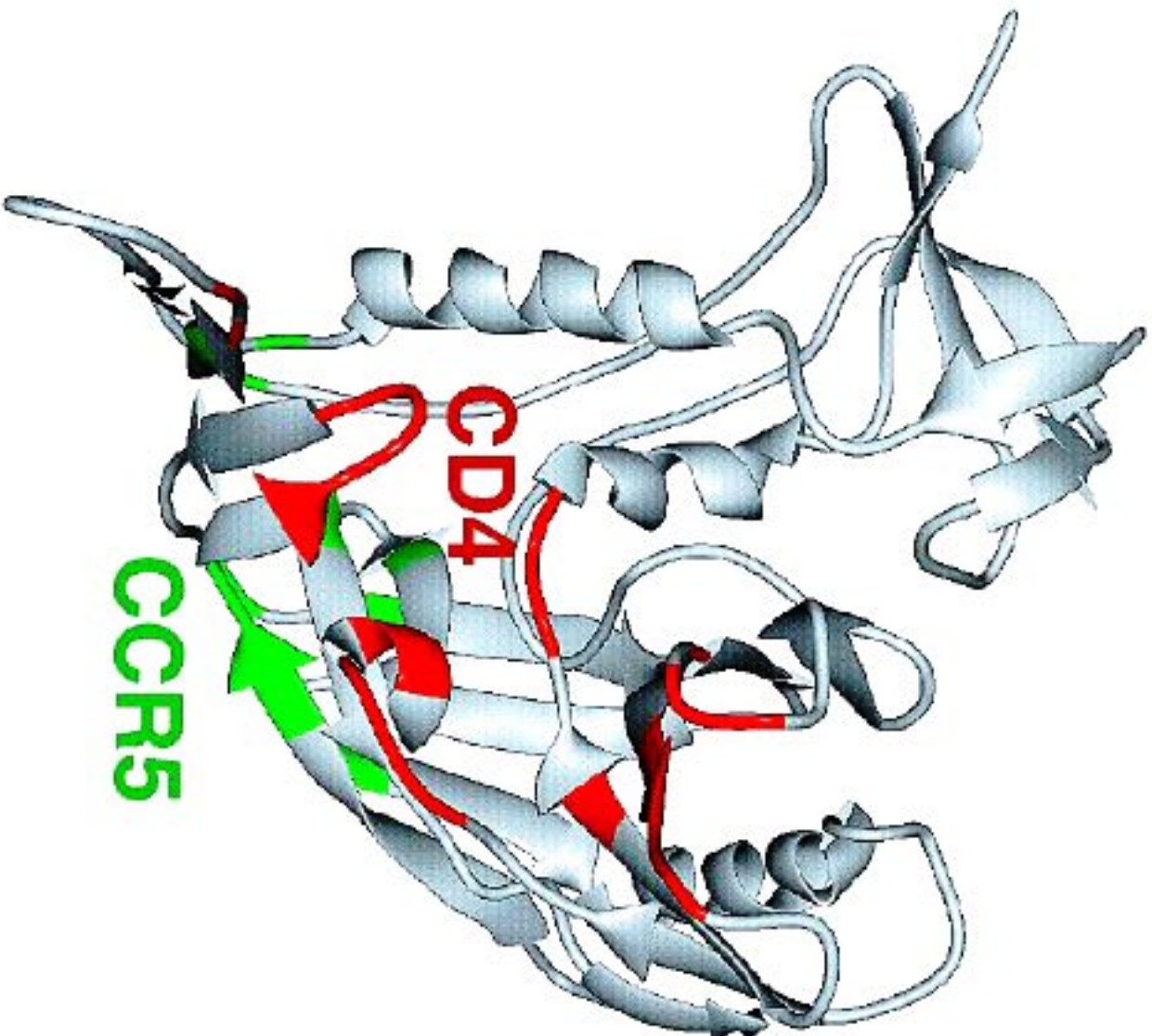


Figure 2c. gp120 core; residues (red) within 4 Angstroms of CD4, important binding residues (green).

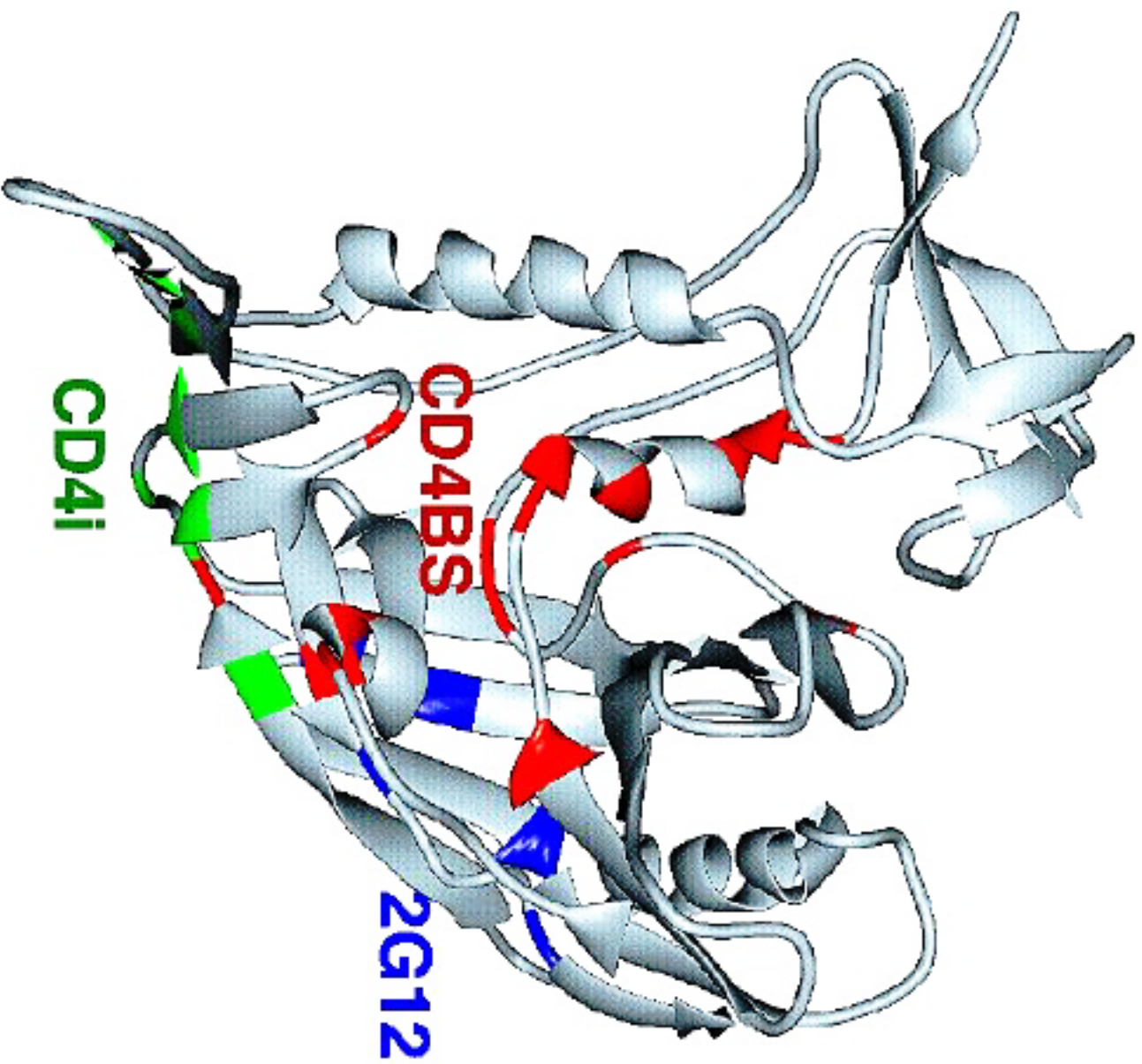


Figure 2d. gp120 core; residues implicated in binding neutralizing antibodies.