Structural Alterations of a Vaccine Target: Clade-specific Differences and Immune Escape of HIV-1 Surface Protein

S. Gnanakaran, T-6; Cynthia Derdeyn, Emory University; Bette Korber, T-6

The genetic diversity of HIV-1 is characterized by a relatively small number of genetically defined clades, or subtypes, A to K, and their recombinants. The envelope (Env) glycoproteins, gp120 and gp41, are the main targets of antibody (Ab) neutralization and are among the most variable of HIV proteins, with typical inter- and intra-clade differences of 20-35% and 10-15% respectively. An antibody-based HIV vaccine would ideally be capable of neutralizing viruses from diverse variants. Whether this will be feasible, and how one might design a polyvalent cocktail that could improve the crossreactive breadth of vaccine-induced responses, can be informed by detailed examination of clade-specific differences in structure and mutational patterns.

Codon-specific ratios of nonsynonymous to synonymous substitution rates (dN/dS) are dramatically different in the B and C subtypes in C3 and V3 regions of gp120. Subtype C is more variable in the C3 region, which is relatively conserved in the B subtype. Conversely, the V3 domain of subtype C exhibits less sequence variation compared with subtype B. Such differences could result from the evolution of lineage-specific structural or functional constraints in the proteins. They could also be due to transmission pressures, or spatially localized differences in neutralizing antibody binding sites, or different HLA frequencies in the circulating populations and the consequent immune escape pressures. By utilizing a synergetic approach that combines sequence and structure, we explore mutational patterns and their structural implications to better understand how positive selection might be driven by immune escape.

We have utilized a variety of simulation techniques to characterize the structure, motion, and thermodynamics of gp120 of B and C clades. The characterization includes longtime all-atom molecular dynamics simulations to capture the dynamics of the gp120 protein, replica exchange enhanced sampling method to capture the conformational variability of loops, and coarse-grained models to capture interactions of V3 with the gp120 core. Coupling these simulation studies with phylogenetic analysis provided several pieces of evidence of sequence and structure differences in viral glycoprotein gp120 between clades B and C. In the C clade, the C3 region α2-helix exhibits high sequence entropy at the polar face but maintains its

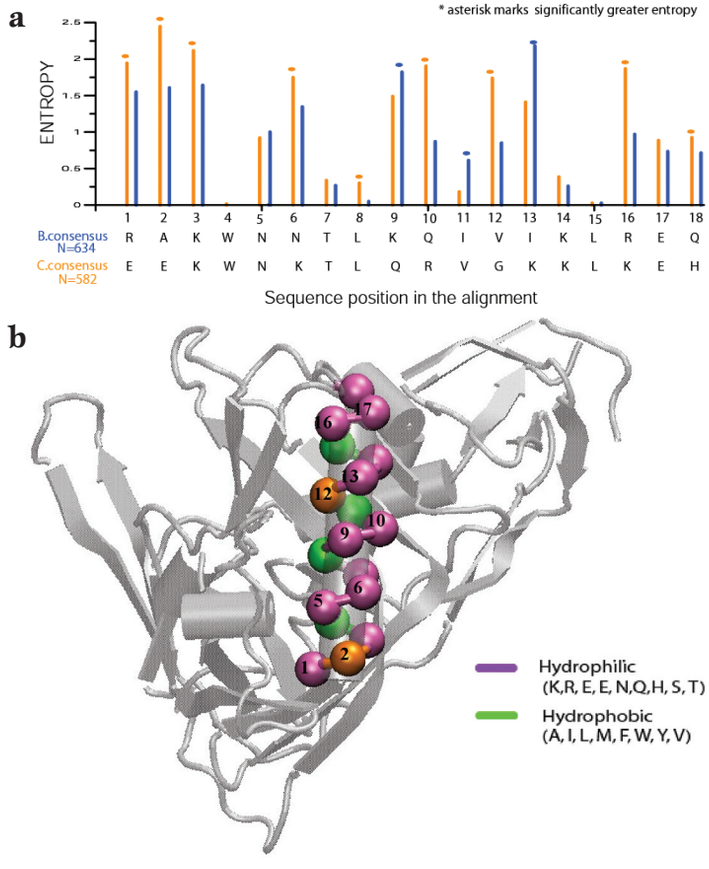

Fig. 1. (A) Sequence entropy profiles of α2 helix for B (blue) and C (orange) clades. (B) Mapping of residue positions from (A) onto the X-ray structure of gp120.

amphipathicity, whereas in the B clade it accommodates hydrophobic residues (Fig. 1). The V4 hypervariable domain in the C clade is shorter than that in the B clade. Generally, shorter V4 loops are incompatible with a glycine occurring in the middle of the α2-helix in the C clade, an intriguing association that could be exploited to inform the design of Env immunogens.

Recently, we have been investigating why the C clade V3 domain lacks sites of strong positive selection analogous to those found in the B clade V3. In the all-atom simulation, we observe the formation of a cluster of hydrophobic residues flanking the V3 tip of C clade (Fig. 2). Importantly, results from the clinical studies at Emory Vaccine Center are consistent with the existence of such a hydrophobic cluster. Stabilizing forces may drive this hydrophobic cluster to avoid solvent exposure by packing within V3 or, possibly, into the core. Our coarse-grained models provide potential sites in gp120 core that interact with this hydrophobic cluster. It is likely that V3 may not be exposed in the C clade since it does not serve as an antibody-mediated neutralization target. Our studies indicate that sequence conservation preserved in structurally specific places within V3 of the C clade may be enabling the V3 itself to play a role in concealing its epitopes.

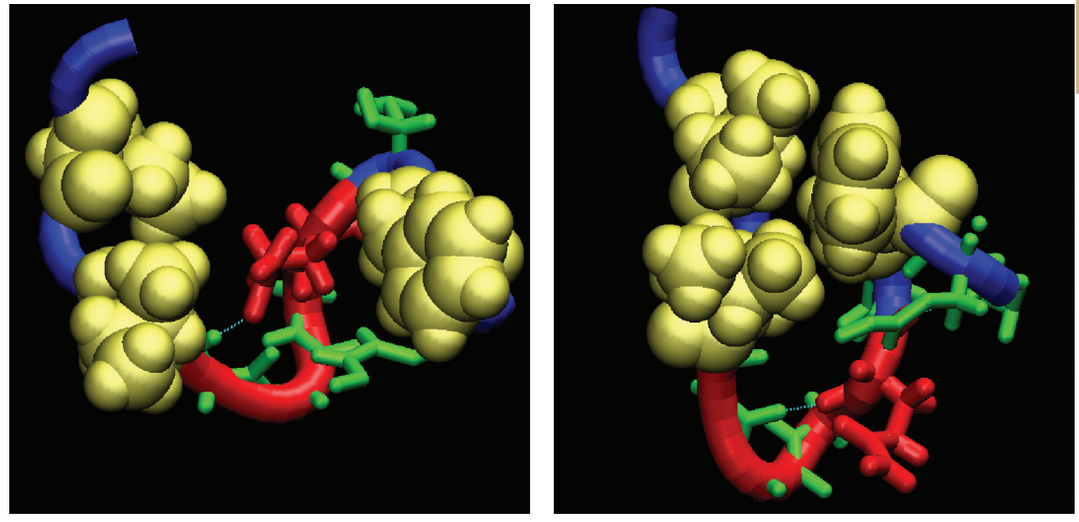

Fig. 2. Formation of a hydrophobic cluster (anchored by residues I307, I309, and F317) near the V3 loop tip.

For further information contact S. Gnanakaran at gnana@lanl.gov.

References

S. Gnanakaran, et al., *J. Virol.* **81,** 4886-4891 (2007).
S. Rong, et al. *J. Virol.* **81,** 5658-5668 (2007).
R.M Lynch et al., LA-UR-08-05424 (2008).
R.M Lynch et al., Aids Res Hum Retroviruses, in press (2009).

Funding Acknowledgments
LANL Directed Research and Development Program