Synopsis of the Thesis

TITLE: HIV-1 immunogen design:

Envelope protein minimization, stabilization and glycan removal

By: Ujjwal Rathore

Thesis Supervisor: Prof. Raghavan Varadarajan

Molecular Biophysics Unit

Indian Institute of Science, Bangalore-560012, India.

Effective vaccines exist for a number of viral diseases. However, despite much effort, there is no successful vaccine against the Human Immunodeficiency Virus (HIV-1). The envelope glycoprotein (Env) of HIV-1 gp120, the most exposed protein on the viral surface is an important candidate for vaccine design. The primary reason for the difficulty in generating an effective neutralizing antibody response against HIV-1 lies in extensive Env sequence variability. Apart from sequence variation, HIV-1 deploys a number of defence mechanisms to evade the host immune response. The gp120 protein is one of the most heavily glycosylated viral proteins known with ~50% of its molecular mass contributed by glycans. The extensive glycosylation of gp120 helps to mask critical neutralization epitopes, although some bNAbs target glycan dependent epitopes in order to achieve virus neutralization. In addition, the presence of long, variable immunodominant loops focuses the immune response away from conserved epitopes. Certain conserved epitopes are cryptic and are exposed only after gp120 binds to CD4. Moreover, the highly flexible nature of gp120 and the labile nature of the gp120:gp41 complex makes structural characterization extremely difficult. Shed gp120 samples various non-native conformations that are likely to be absent in native Env and that elicit

non-neutralizing antibodies. Despite these defence mechanisms, about 20-30% of HIV-1 patients do generate a broad neutralization response and around 1% of them do so with high potency. Although these bNAbs and their epitopes have been identified, eliciting similar bNAbs through immunization is challenging.

Monomeric gp120 is an obvious candidate immunogen, due to its surface accessibility and immunogenicity during natural infections. Initial immunization studies with full length monomeric gp120 failed to elicit neutralizing antibodies and failed to show protection in a human clinical trial. This indicates that neutralization epitopes are not optimally presented on monomeric gp120. Elicitation of neutralizing antibodies is a desirable feature of a vaccine. Thus a rational design approach to focus the immune response towards specific epitopes targeted by known bNAbs is urgently required. The aim of any rational immunogen design methodology targeting HIV-1 is to maximize the exposure of conserved neutralization epitopes and to simultaneously ensure minimal exposure of variable non neutralizing epitopes to the immune system. This can likely be achieved either by (a) stabilization of native Env trimers, or/and by (b) protein fragment design.

Chapter 1 gives a brief overview of the HIV-1 virus, and outlines the structural organization of its Env protein. The difficulties in generating an effective HIV-1 vaccine and the recent advances in rational structure based HIV-1 vaccine design are discussed.

The outer domain (OD) of HIV-1 Env glycoprotein gp120 contains the site for the binding of cellular receptor CD4 and epitopes for a large number of recently discovered broadly neutralizing antibodies (bNAbs). Therefore, OD is considered to be an important candidate for structure-based vaccine design and for designing minimal gp120 immunogens. We have previously reported the design and

characterization of a non-glycosylated, *E. coli* expressed outer domain immunogen (OD_{EC}) that bound CD4 and bNAb b12 with micromolar affinity and elicited a modest neutralizing antibody response in rabbits (Bhattacharyya et al., 2010a). In **Chapter 2**, we report the design and characterization of an engineered bridging sheet deleted variant of OD_{EC} ($OD_{EC}\Delta BS$), which showed significant improvement in biophysical properties over the original OD_{EC} molecule (ΔT_m = 6.6°C) and bound bNAb VRC01 with improved affinity (K_D of ~100-150nM). In another approach, we engineered an OD molecule (OD_{EC}Consensus) using consensus based protein design methodology. Consensus mutations are generally believed to improve the thermal stability of proteins. OD_{EC}Consensus showed improvement in biophysical properties (ΔT_m = ~3.0°C) and in binding affinities with bNAbs such as VRC01 and b12 (K_D of ~60-80nM) as compared to OD_{EC}. When these two improved designs were used as immunogens in rabbits, the resultant sera showed significantly higher anti-gp120 titers as compared to the original OD_{EC} design and in addition, showed cross-clade neutralization of Tier 1 viruses.

HIV-1 Env protein gp120 is heavily glycosylated with approximately 25 glycosylation sites of which ~4 are located in the inner domain, ~7-8 in the V1V2 and V3 variable loops and the rest in the outer domain (OD) of gp120. Glycans partially shield Env from recognition by the host immune system and are believed to be indispensable for proper folding of gp120. It is not known how many of the ~25 glycans present on the gp120 surface are essential for folding. This is an important question as these glycans mask various conserved epitopes on the viral surface and thereby impede the development of an effective gp120 based vaccine. In addition, several recently isolated bNAbs bind to specific glycans on gp120. If the importance of each PNGS in maintaining the structural and functional integrity of gp120 could

be determined, then this information could be used to modulate coverage of Env by glycans.

In chapter 3, Using rationally designed mutations and yeast surface display (YSD), we show that glycosylation is not essential for the correct *in vivo* folding of gp120 outer domain fragment alone or outer domain in the context of core gp120. Following randomization of the remaining four glycosylation sites and screening by yeast surface display, two core gp120 mutants were isolated which contained a single inner domain glycan, retained yeast surface expression and bNAb binding. This demonstrates that most core gp120 glycans are dispensable for folding in the absence of gp41. Presence of either N88 or N241 glycan was found to be essential for the folding of core gp120. Both N88 and N241 belong to the inner domain of gp120, which interacts with gp41. Since gp41 was absent in the core gp120 constructs, we hypothesized that removal of these inner domain glycans might result in exposure of hydrophobic surface and consequently aggregation of core gp120. An OD_{EC} variant lacking all 14 potential N-glycosylation sites (PNGS) was purified from *E.coli* (ΔG-OD_{EC}). ΔG-OD_{EC} is devoid of the inner domain, part of the bridging sheet, V1/V2/V3 variable loops and has a fivefold lower molecular weight than full length gp120. However, it could still bind bNAb b12, bNAb VRC01 and CD4 with affinities comparable to that of full length gp120 (K_D of ~10-50nM).

Human bNAbs that prevent viral entry into cells arise after several years of infection, but are extensively mutated. Unmutated (germline) versions of these antibodies generally do not show measurable binding with WT gp120. Hence such antibodies are difficult to elicit through immunization with Env. It has been proposed that immunogens that bind well to germline precursors of bNAbs as well as the corresponding mature antibodies may help to elicit such bNAbs upon

vaccination. Glycan free ΔG-OD_{EC} could bind to the germline variant of bNAb VRC01 with high affinity (K_D of ~10nM) without any requirement for additional mutations. The results from Chapter 3 inform immunogen design, targeting mature as well as GL bNAbs, clarify the role of glycosylation in gp120 folding and illustrate general methodology for design of glycan free, folded protein derivatives.

Apart from gp120 folding, glycosylation is also considered to be essential for Env native trimer formation and for viral infectivity. Most of the studies done to examine the effect of glycan removal on viral infectivity were performed with single site mutants. Whenever multiple sites were mutated in combination, it resulted in loss of viral infectivity which led to the hypothesis that a certain level of glycan coverage is essential for maintaining infectivity of the virions. In **Chapter 4**, we studied the importance of glycosylation in maintenance of native Env trimers and viral infectivity. By introducing rational mutations at potential N-glycosylation sites, we could show that Env lacking all core gp120 glycans can retain native trimeric structure. We used pseudoviral neutralization assays to show that infectivity can be retained in the absence of all glycans from core gp120, if mutations are introduced rationally. Loss of infectivity, in the earlier studies can be attributed to non-optimal substitutions (hydrophilic and uncharged mutations) at the PNGS, which probably failed to rescue the folding of Env, in the absence of glycans.

We also found that recognition of a germline reverted version of bNAb VRC01 increases substantially with the progressive loss of glycans from JRFL pseudoviruses, indicating that one way by which glycans can shield the virus from the host immune response could be by reducing the interaction with germline antibodies. These deglycosylated constructs can also be used in DNA immunizations as they will quite likely activate a large number of germline variants including those

of bNAbs. Combinations of these molecules in various prime: boost regimes will be helpful in expanding our understanding about the role of glycosylation in modulating the gp120 directed immune response. Knowledge of these mutations which can be tolerated at PNGS in gp120 would also be useful for structure determination of full length gp120 and native Env trimers in the future as extensive glycosylation has rendered this task extremely difficult. Further, the methodology outlined in this work can in principle be used to probe the role of glycans in the stability and folding of any glycosylated protein.

In **Chapter 5**, we have attempted to further improve the binding and biophysical properties of OD_{EC} immunogens by reducing their conformational flexibility. As disulfides have long been known to stabilize proteins by reducing the entropy of the unfolded state, we attempted to rationally engineer cross-strand disulfides in $OD_{EC,}$ in order to reduce the conformational flexibility of the molecule. In another design, we have introduced previously known cavity filling mutations to reduce the flexibility of OD_{EC} by restricting it in the 'CD4 bound state'. Finally, we combined the rational stabilizing mutations with the glycosylation site mutations to generate second generation OD immunogens capable of binding to various CD4 binding site (CD4bs) ligands with affinities comparable to that of full length WT gp120 (~10-50nM K_D). In another slightly different approach to design an OD immunogen, we made a cyclic permutant of OD_{EC} (CycV4OD), which resulted in a drastic improvement in the binding affinities for CD4, bNAb b12 and bNAb VRC01 (~10-45nM K_D) without any additional mutations. When tested in rabbits, the two immunogens designed in the current study, showed improvement in their ability to generate anti-gp120 sera and neutralizing antibody against Tier 1 viruses relative to the previous immunogens.