

## Chapter 39

# Structural Biology and the Design of Effective Vaccines for HIV-1 and Other Viruses

Peter D. Kwong, Gary J. Nabel, Priyamvada Acharya, Jeffrey C. Boyington, Lei Chen, Chantelle Hood, Albert Kim, Leopold Kong, Young Do Kwon, Shahzad Majeed, Jason McLellan, Gilad Ofek, Marie Pancera, Mallika Sastry, Anita Changela, Jonathan Stuckey, and Tongqing Zhou

Structural biology provides a wealth of information about the three-dimensional organization and chemical makeup of proteins. An understanding of atomic-level structure offers enormous potential to design rationally proteins that stimulate specific immune responses. Yet current vaccine development efforts makes little use of structural information. At the Vaccine Research Center, a major goal is to apply structural techniques to vaccine design for challenging pathogens, that include human immunodeficiency virus type 1 (HIV-1) and other enveloped viruses such as influenza, Ebola, and respiratory syncytial viruses. Our three-part strategy involves 1.) the definition of the functional viral spike at the atomic level 2.) achieving a structural understanding of how neutralizing antibodies recognize the spike, and 3.) rational development of proteins that can elicit a specific antibody response. Overall, our strategy aims to incorporate information about viral spike-antibody interactions, to assimilate immunogenic feedback, and to leverage recent advances in immunofocusing and computational biology.

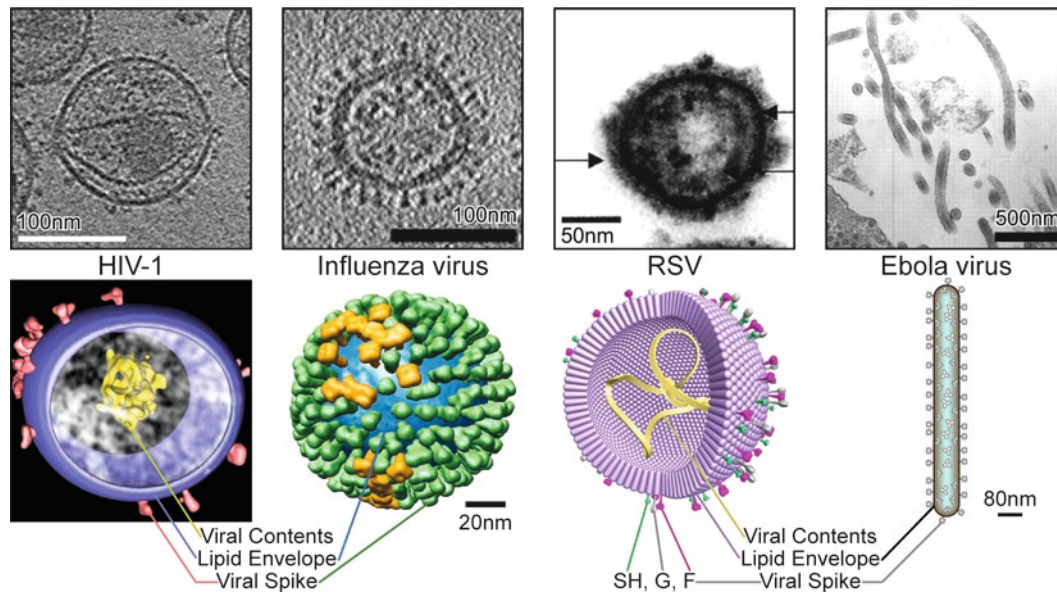
For an elusive pathogen such as HIV-1, whether such a strategy will succeed depends both on the existence of sites of envelope (Env) vulnerability susceptible to neutralizing antibody and on the ability of the human immune system to generate high titers of such antibodies. With HIV-1, the atomic-level definition of one such vulnerable site [1] and the discovery of individuals with broadly neutralizing sera that target this site [2] bodes well for our informatics-based approach. With Ebola and influenza viruses, structures of functional viral spikes have been determined [3, 4], and sites of antibody vulnerability and elicitation of appropriate antibodies are under investigation. With respiratory syncytial virus, structures of closely related spikes have been determined [5], and issues center on the elicitation of antibodies that defuse rather than exacerbate disease.

### 39.1 Enveloped Viruses and Neutralizing Antibodies

A hallmark of enveloped viruses is the presence of a lipid membrane that surrounds the viral core structural proteins (Fig. 39.1). These membranes are derived from the host cell, and among other purposes, serve to protect the internal components of the virus from immune recognition. While protective, the membrane is also a barrier: to infect a cell, enveloped viruses must breach not only the target host membrane, but also their own membranes. To do so, they utilize energy stored in metastable viral spikes that protrude through the viral membrane [6]. These spikes also target the virus to specific receptors on host cells, prior to the fusion event. Together, these processes facilitate entry of the viral genome to the host cell.

In the case of type 1 fusion machinery (Fig. 39. 1), the viral spike is composed of a trimer of glycoproteins, each with a single transmembrane domain. The glycoproteins are cleaved during the maturation process into exterior N-terminal and transmembrane C-terminal components. The N-terminal component typically interacts with appropriate host receptors. The C-terminal component is involved in membrane fusion, which it accomplishes by inserting a hydrophobic “fusion peptide” into the target cell membrane, and then refolding to bring fusion peptide and transmembrane region/viral membrane into close proximity [7].

Because it protrudes through the protective membrane, the viral spike is a target for neutralizing antibody, which either binds to the spike to prevent cell attachment or binds and prevents conformational changes required for fusion. Vaccine efforts to elicit neutralizing antibodies that target the viral spike are currently underway for a number of



**Fig. 39.1** Enveloped viruses with type 1 fusion machinery. Cryo-electron tomograms or electron micrographs (top row) are shown for HIV-1, influenza virus, RSV and Ebola virus. Schematics (bottom row) point out locations of viral spike, lipid envelope, and viral contents. HIV-1 cryo-electron tomogram and rendering reproduced with permission from the work of Sriram Subramaniam and colleagues, NCI/NIH (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109–113. Influenza cryo-electron tomogram and rendering reproduced with permission from Harris, A., Cardone, G., Winkler, D.C., Heymann, J.B., Brecher, M., White, J.M. and Steven, A.C. (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci USA* 103:19123–19127, copyright (2006) National Academy of Sciences, U.S.A.

RSV electron micrograph reproduced with permission from Brown G, Aitken J, Rixon HW, Sugrue RJ. (2002) Caveolin-1 is incorporated into mature respiratory syncytial virus particles during virus assembly on the surface of virus-infected cells. *J Gen Virol.* 83(Pt 3):611–21, copyright (2002) Society for General Microbiology. Ebola electron micrograph reproduced with permission from Halfmann, P., Kim, J.H., Ebihara, H., Noda, T., Neumann, G., Feldmann, H. and Kawaoka, Y. (2008) Generation of biologically contained Ebola viruses. *Proc Natl Acad Sci USA* 105:1129–1133, copyright (2008) National Academy of Sciences, U.S.A.

human viral pathogens. Here we describe efforts with four enveloped viruses, all of which utilize type 1 fusion entry machines.

## 39.2 HIV-1

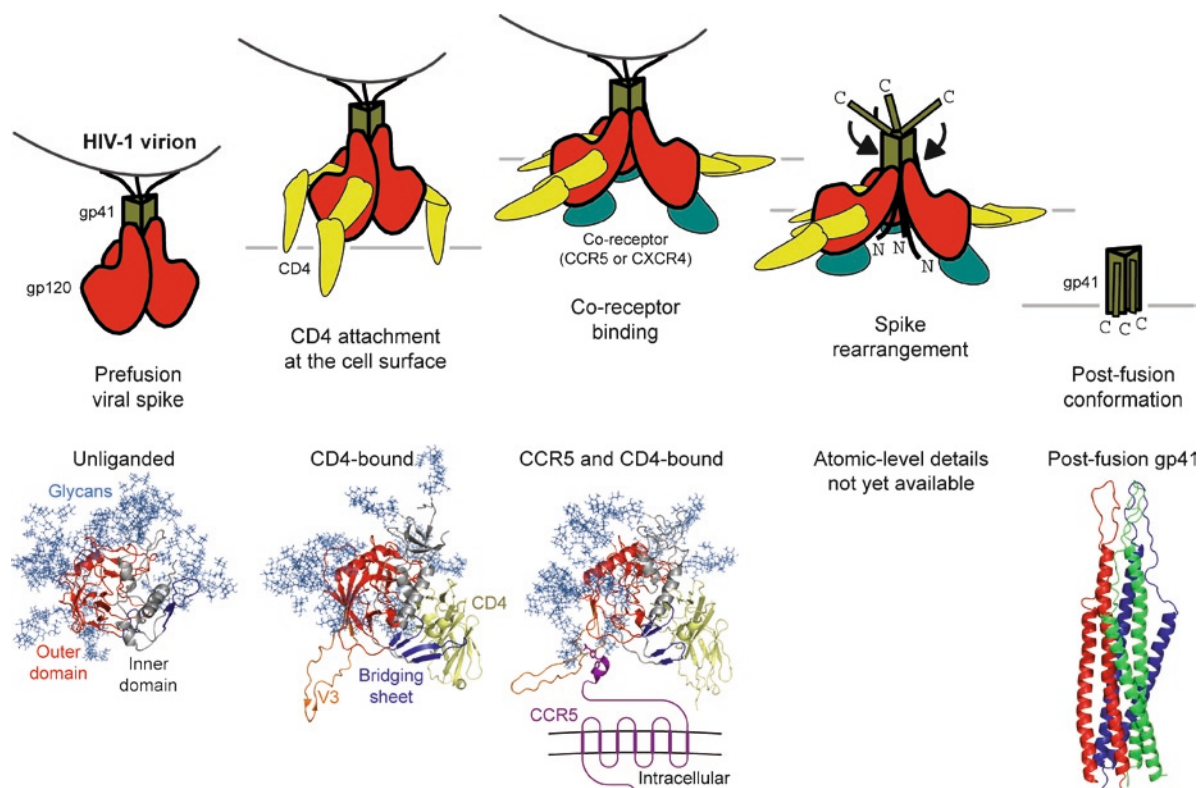
HIV-1 is an enveloped lentivirus. Since its crossover from chimpanzees to humans at the end of the 19<sup>th</sup> or early part of the 20<sup>th</sup> centuries [8], the human immunodeficiency virus has killed more than 20 million individuals, and an additional ~30–35 million are currently infected [9].

Significant effort has been expended to develop an effective HIV-1 vaccine. To date, however, no vaccine against HIV-1 has been developed that elicits a fully protective response in humans. In large measure, this lack of progress is related to the elusive properties of the HIV-1 Env, which is highly variable, silenced by carbohydrates, and flexible in conformation [10]. We have therefore sought to understand the structure of the HIV-1 Env in an effort to reveal vulnerabilities on the Env glycoproteins, gp120 and gp41.

Three gp120 envelope glycoproteins associate non-covalently with three gp41 transmembrane molecules to form the functional viral spike (reviewed in [11]). Despite extensive

effort, by our group and others, the trimeric functional spike has so far resisted atomic-level structure determination. Recent low resolution cryo-electron microscopy (EM) studies have provided insight into possible gp120-gp41 subunit arrangements within a functional viral spike [12, 13]. Information from cryo-EM studies is being used in our efforts to design candidate Env molecules amenable to high resolution crystallographic analysis. Part of the difficulty in structurally characterizing the functional spike relates to the biology of HIV-1. Both gp120 and gp41 components of the viral spike undergo conformational rearrangements during viral entry (Fig. 39.2). The binding by the gp120 component to the CD4 receptor is believed to induce large structural rearrangements [14, 15], which are required for the formation of the binding site for a requisite coreceptor [16], generally CXCR4 or CCR5 (reviewed in [17, 18]). The gp41 component then utilizes a large conformational change, involving formation of a trimeric “post-fusion” coiled-coil [19–23], to fuse the viral and target cell membranes. Since the “pre-fusion” conformation of the viral spike [24] is believed to be the primary target for elicitation of neutralizing antibody, we have sought to understand its conformation in atomic-level detail.

Our current approaches focus on developing strategies to overcome the inherent flexibility of the envelope spike and to



**Fig. 39.2** HIV-1 entry pathway and atomic-level structures. (Top row) Schematics of the viral spike (gp120 in red and gp41 in tan) are shown, first in the context of free HIV-1 virion (left-most image), then binding to CD4 (yellow) at the target cell surface. Binding to CD4 induces conformational changes that assemble the binding site for a requisite co-receptor (teal). Binding to co-receptor results in further conformational changes: the fusion peptide at the N-terminus of gp41 is thrown into the target cell membrane; the gp41 C-terminus rearranges to form the

coiled-coil structure with gp41 N and C termini in close proximity (right-most image). (Bottom row) Atomic-level structure of monomeric gp120 and trimeric gp41 are shown in  $\alpha$ -ribbon representation. Structures are arranged to correspond with conformational states shown in the schematics. Unliganded gp120 coordinates from (15); PDB ID 2bf1. CD4-bound gp120 coordinates from (29); PDB ID 2b4c. CCR5- and CD4-bound gp120 coordinates from (30); PDB IDs 2qad & 2rll. Post-fusion gp41 coordinates from (19); PDB ID 2ezo

isolate sufficient amounts of conformationally stable trimeric proteins for high resolution structural studies. In an effort to obtain well-diffracting crystals of the “pre-fusion” trimeric viral spike, we have utilized heterologous trimerization domains as well as ligands and chemical/mutational modifications to enhance protein stability. Most of the viral spikes that we have constructed have mutations that prevent the proteolytic cleavage of gp41 from gp120, which increases protein stability and ease of production. However, since cleaved viral spikes are antigenically and structurally distinct from uncleaved spikes, and represent the functional form of the molecule, methods for creating cleaved viral spikes are being optimized. gp41 peptides are currently being developed that bind to the N and C termini of gp120 and the co-crystallization of other published peptides is also being pursued [25].

In a further attempt to increase the number of crystallization variants [26], the methods described above are being applied to envelope proteins from the closely related primate lentiviruses, HIV-2 and SIV. HIV-1, HIV-2 and SIV envelope glycoproteins share a high degree of sequence conservation and common structural elements. Like HIV-1, HIV-2 and

SIV use CD4 and chemokine co-receptors, thereby suggesting the use of a common mechanism for viral entry. High resolution structure determination of a trimeric HIV-2 or SIV envelope spike could provide a useful prototype for HIV-1 immunogen design.

We and others have obtained structural information on the individual gp120 and gp41 components (Fig. 39.2). For gp41, only post-fusion structures have been determined [19, 21–23]. For gp120, structures of several proposed intermediates in the entry pathway have been determined. These structures include a flexible, unliganded gp120-core for a simian immunodeficiency virus isolate [15], the CD4-bound gp120 core for several different HIV-1 isolates [27–29], and HIV-1 gp120 bound to both CD4 and the N-terminus of CCR5 that was obtained by a combination of X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and molecular docking [30]. These structures reveal the conformational flexibility that gp120 must adopt to facilitate HIV-1 entry.

In addition to the entry mechanism, the structural information reveals additional mechanisms of humoral immune evasion. Mapping antibody epitopes onto atomic-level structures

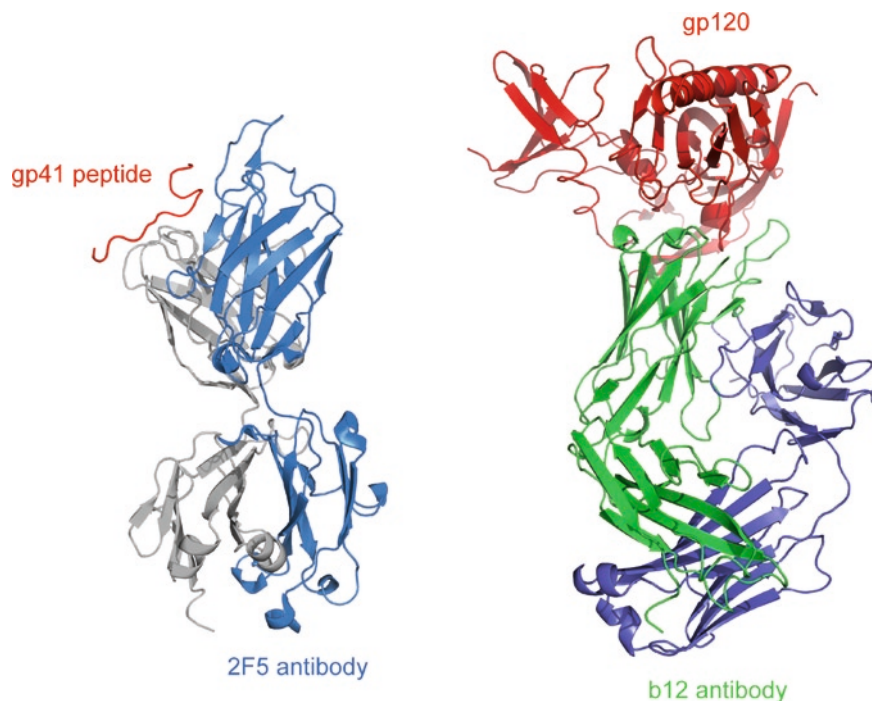
showed that a large portion of the gp120 surface is not recognized by antibody [10]. This surface corresponds to regions with high concentrations of N-linked carbohydrate (glycan), suggesting that carbohydrate is seen as “self” by the humoral immune system and can form an immunologically “silent” face [10]. In the context of the functional viral spike, this glycan can protect neighboring epitopes through an “evolving glycan shield” [31]. These and other studies, defining for example protective variable loops [32] and conformational masking of antibody epitopes [33], have led to an understanding of the molecular mechanisms that protects HIV-1 from the humoral immune response.

### 39.3 Structure-Assisted Vaccine Design for HIV-1

Currently, two lines of investigation have shown promise in our effort to harness structural biology for vaccine design. One line involves the delineation of functional constraints to identify potential footholds of conservation and exposure. We have investigated antibodies, called CD4-induced antibodies that bind to the co-receptor-binding site on gp120. These were found capable of neutralizing not only HIV-1, but also

the more evolutionarily divergent HIV-2 [34]. They also accumulate to high titers in most HIV-1 infected individuals [34]. Unfortunately, the virus effectively hides the site of co-receptor binding prior to CD4 engagement [34, 35]. After CD4 engagement at the cell surface, the close proximity of the target-cell membrane sterically occludes potential sites of vulnerability on the co-receptor-binding site [36]. These studies demonstrate that functional constraints restrict epitope variation but do not necessarily expose epitopes sensitive to antibody neutralization. We are currently exploring whether the functional constraint of binding to the CD4 receptor, which — unlike the co-receptor-binding site — must be available as an initial site of attachment [1], can provide a site of vulnerability to neutralizing antibodies.

A second line of investigation involves structural characterization of monoclonal antibodies that neutralize diverse isolates of primary HIV-1 and comparative analysis to antibodies that do not neutralize the primary isolates. Only four antibodies that neutralize the primary isolates have thus far been characterized—the monoclonal antibodies 2F5, 2G12, 4E10, and b12 [37–40]. We have determined the structures of 2F5 and b12, each with their HIV-1 envelope epitopes [1, 41] (Fig. 39.3). In the case of the broadly neutralizing 2F5 antibody, the structure of the membrane-proximal external region of gp41, when bound by 2F5, was found to adopt an extended



**Fig. 39.3** Structure of antibodies 2F5 and b12 in complex with their respective HIV-1 Env epitopes. (Left) 2F5 antibody with gp41 epitope. The antigen-binding portion (Fab) of 2F5 is shown with heavy chain in blue and light chain in gray. A peptide, corresponding to the gp41-linear epitope

recognized by 2F5, is shown in red. Coordinates from (41); PDB ID 1tji. (Right) b12 antibody with gp120. The Fab of b12 is shown with heavy chain in green grasping its gp120 epitope, while the light chain in dark blue is positioned over 15 Å away from gp120. Coordinates from (1); PDB ID 2ny7

loop conformation – distinct from unbound structures of this region, which largely adopt alpha-helical conformations [41]. These structural differences prompted the hypothesis that the inability of existing immunogens against this region to elicit 2F5-like antibodies may partly be due to a lack of proper antigenic mimicry of the 2F5-bound form of gp41. In collaboration with Drs. David Baker and Bill Schief at the University of Washington, techniques of computational protein design are being used to develop heterologous epitope scaffolds that possess conformational mimics of the 2F5-bound form of gp41 on their surfaces. These epitope scaffolds not only provide a means for structural mimicry, but also serve to focus the immune response to specific sites within otherwise complex structures.

The broadly neutralizing antibody, IgG1 b12, primarily recognizes the conformationally invariant outer domain of gp120. Focusing on the conserved CD4-binding loop, the b12 epitope overlaps considerably with the CD4 contact surface on gp120 and extends further towards the glycan covered “silent face”; however, b12 has only peripheral interactions with the other conformationally mobile parts of the gp120, such as the residues that form the bridging sheet in the CD4-bound state of gp120. Our research indicated that b12 exploits the conformationally invariant and functionally conserved CD4-contact site to achieve neutralization [1]. With a vulnerable site on HIV-1 gp120 defined, we are pursuing collaborative studies to generate b12-binding immunogens in a number of different protein forms. These alternative forms include trimeric mimics of the viral spike, monomeric variants of gp120, domain versions of gp120 that retain b12 binding, and epitope-scaffold mimics believed to be unencumbered by most Env-based mechanisms of humoral evasion. Tests of these various formats in small animals should reveal their potential to elicit antibodies similar to the template broadly neutralizing ones.

In addition to analysis of the b12 antibody, it may be useful to also analyze antibodies that target the CD4-binding site (CD4BS), but do not effectively neutralize HIV-1. Such antibodies comprise a high percentage of the elicited antibody response in HIV-1 infected individuals. We have succeeded in determining the structure of F105 in complex with gp120 (Lei Chen and Young Do Kwon, personal communication). F105 is a prototypical CD4BS antibody and has been tested in Phase I clinical trials [42]: it shows broad recognition of monomeric gp120, competes with CD4 for binding, and neutralizes laboratory-adapted – but not primary – isolates of HIV-1. Analysis of the epitopes on gp120 for F105, b12 and CD4 shows they all bind to a very closely related site on gp120. Despite their similarity, one large difference we find is that the region that makes up the bridging sheet in the CD4-bound conformation of gp120 opens up in the F105-gp120 structure. Strands  $\beta$ 20/ $\beta$ 21 open to uncover

a conserved hydrophobic surface, which serves as a focus of F105 binding. This suggests that antibodies like F105 need to access the hydrophobic region under the bridging sheet. To define F105 as a prototype for non-neutralizing CD4BS antibodies, we tethered the  $\beta$ 20/ $\beta$ 21 strands to the inner domain with a disulfide and observed inhibition of binding for most CD4BS antibodies, though not b12 or CD4. Thus, this result highlights the importance of occluding the hydrophobic region under the bridging sheet in vaccine immunogens designed to elicit b12-like antibodies, and also provides another example for how atomic-level structural information can be used in the development of a successful HIV vaccine.

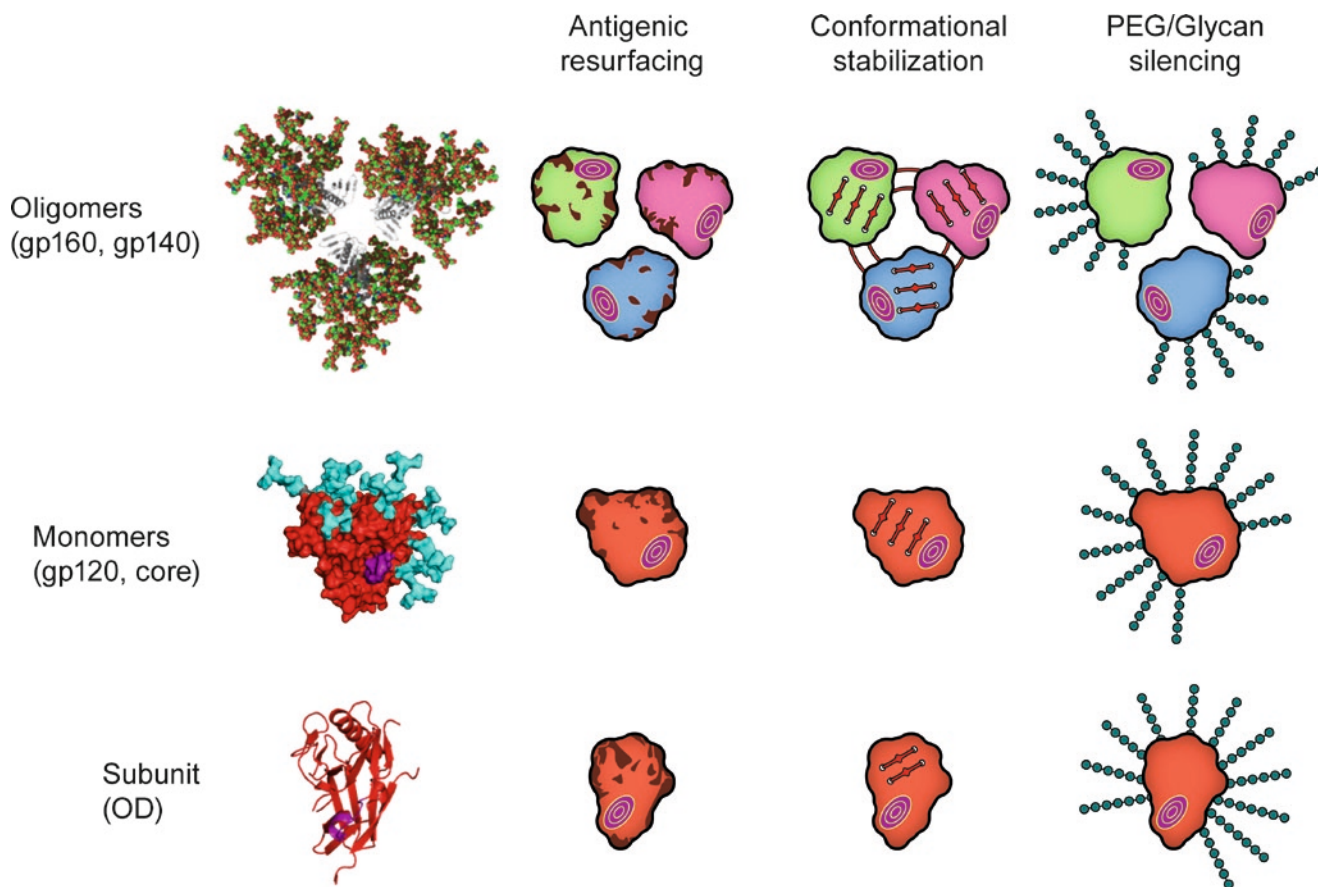
In addition to our efforts to understand how HIV-1 can effectively be neutralized by CD4BS antibodies, we are also studying small molecules and CD4 mimetic miniproteins that inhibit the binding between gp120 and CD4. Small molecules such as BMS806 potently inhibit HIV at nanomolar concentrations [43]. Others such as NBD-556 and related compounds do so at micromolar concentrations [44]. Definition of the atomic-level mechanism by which these small molecules inhibit HIV may provide insight into how antibodies might similarly neutralize.

Collaborative efforts with Dr. Joseph Sodroski at Harvard Medical School to define the mechanism of BMS806 neutralization are ongoing. Mutations in gp120 in both inner domain and V3 region affect BMS806, a confusing result as these regions are over 50 Å distal from each other. Efforts to crystallize BMS806 are also underway, but have not yet met with success.

With NBD-556 and related compounds, modeling and exploration of structure-activity relationships suggest an interaction with a deep gp120 pocket (the Phe-43 pocket) [45], and efforts to define the interaction crystallographically are on-going. Currently the results suggest that the narrow entrance to the pocket prevents natural amino acids from fully accessing this site of vulnerability.

In terms of the CD4 mimetic miniproteins [46, 47], we have solved the crystal structures of gp120 with CD4M33/CD4M47, synthetic miniproteins into which the gp120 binding surfaces of the CD4 receptor were transplanted, to understand how precisely these miniproteins mimic CD4 for binding to gp120 [48, 49]. Biphenyl moiety of CD4M33 reaches into the deep Phe43 cavity in a way that no side chains of natural amino acids can emulate [49]. Furthermore these structures show that small molecules and miniproteins bind exactly the same gp120 conformation as CD4 binds and suggest that these molecules are useful not only for therapeutic purpose but also for exposing potential vulnerable sites on gp120 for neutralizing antibodies.

Other methods of focusing the antibody response to a precise site required for broad neutralization are being developed. These include antigenic resurfacing, conformational



**Fig. 39.4** Methods of immunofocusing to enhance the elicitation of b12-like antibodies. Three immunogenic formats of the gp120 Env are shown: as oligomers, corresponding to the trimeric state in the viral spike; as monomers, corresponding to the shed state of gp120; and

as a subunit domain. Each of these formats can be modified by antigenic resurfacing, conformational stabilization and/or PEG/glycan silencing, to focus the immune response onto the b12-identified site of vulnerability

stabilization, and PEG/glycan silencing (Fig. 39.4). These methods are in theory independent from each other and could be used combinatorially.

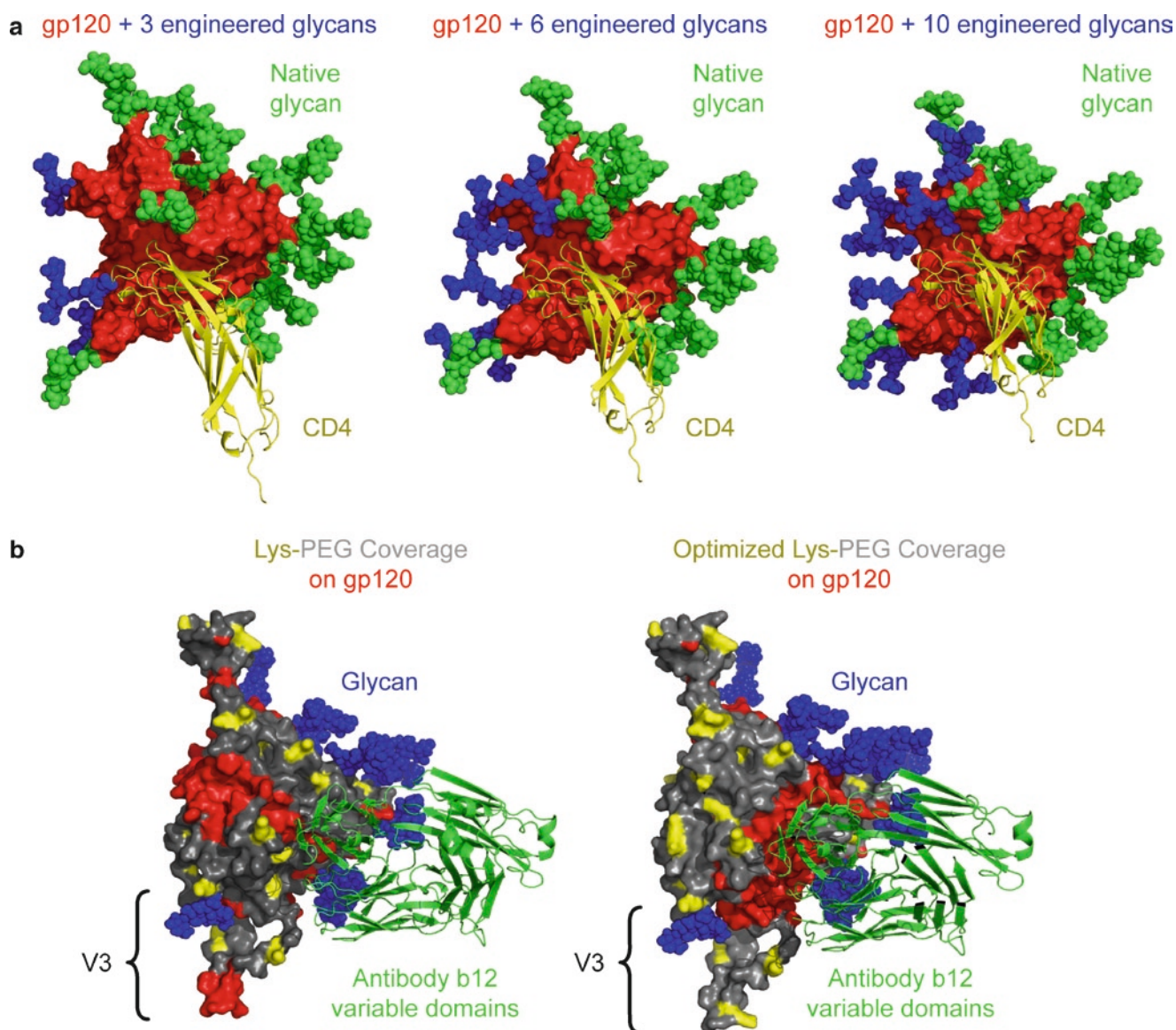
Antigenic resurfacing involves the alteration of all surfaces other than the desired site of recognition. For example, to design antigenically resurfaced immunogens for eliciting b12-like antibodies, the defined b12 epitope was kept in its b12-bound conformation while regions otherwise not covered by glycans were subjected to modification using evolutionary similar sites from SIV or HIV-2. These variants were generated by computational design, expressed and tested for desired antigenic properties, such as binding to b12 and lack of binding to non-neutralizing antibodies. To further promote recognition of this specific region, two immunogens antigenically resurfaced with different amino acid modifications can be used in sequential prime boost immunization combinations.

Conformational stabilization involves the introduction of select mutations that reduce the flexibility of gp120 by locking it into a specific conformation [1]. Conformational stabilization can also increase physical stability, a useful property

in making proteins sufficiently stable to maintain antigenic fidelity in the presence of adjuvants and immunization conditions.

Immunosilencing, meanwhile, seeks to reduce the immunogenicity of regions outside of the target epitope, through the addition of moieties such as N-linked glycan and or polyethylene glycol (PEG), which inherently have low immunogenicity. N-linked glycan, for example, when added to a particular region of a protein appears to dramatically reduce the immunogenicity of that region [50, 51]. As the sequon controlling N-linked glycosylation (Asn-X-Ser/Thr) can be introduced into an immunogen in a structure-specific manner, glycan-silencing allows for targeted immunosilencing (Fig. 39.5a). A related strategy utilizes PEG-chemical modification to generate protein-PEG adducts [52–54]. Because PEG molecules can be specifically coupled to lysine residues, and lysines can be introduced in a structurally specific manner, such PEG-Lys immunosilencing can also be carried out in a structure-specific manner (Fig. 39.5b).

The toolkit of structure-based design is still evolving. Nonetheless, initial attempts in small animals are beginning



**Fig. 39.5** PEG and Glycan silencing of HIV-1 gp120. **(a)** Glycan silencing. The surface of HIV-1 HXBc2 core gp120 is shown in red bound to 2-domain CD4 (yellow;  $\alpha$ -ribbon representation). Coordinates from (28); PDB ID 1gc1. N-linked glycans are shown in all atom representation at three different densities. Endogenous glycans are colored green and engineered glycans in blue. Experiments with Dr. Richard Wyatt's group at the Vaccine Research Center are currently being carried out to assess the effect of these glycan modifications on expression, antigenicity and immunogenicity. **(b)** PEG silencing. The variable portion of the b12 antibody is shown as a

green  $\alpha$ -ribbon. N-linked glycans are colored blue and the surface of gp120 is colored red, except for regions corresponding to Lys residues (colored yellow) or within 10 Å of a lysine (colored gray). The HXBc2 core with V3 is shown in the left image, and the Lys-optimized core is shown in the right image. As can be seen, structure-based optimization allows the b12 binding surface to be free of Lys, and to improve coverage of the rest of the gp120 surface (for example, of the V3 loop). Experiments with Dr. Quentin Sattentau's group at Oxford University are currently being carried out to assess the effect of such targeted PEGylation

to yield results. For example, studies done in collaboration with Dr. Richard Wyatt's group at the Vaccine Research Center have demonstrated the ability of conformational stabilization to focus immune response [55]. When gp120 monomers were stabilized in the CD4-bound conformation with multiple disulfides and used for immunization in rabbits, they elicited a stronger response of neutralizing CD4i antibodies than the native monomer. These studies demonstrate

that structure-directed conformational stabilization can alter the quality and magnitude of the humoral immune response to a desired epitope. Though the CD4i antibodies elicited in this study are not potent against primary isolates, this proof of concept study bolsters our confidence that structure guided immunogen modification which leads to exposure and stabilization of vulnerable sites on the HIV envelope, elicits high titers of antibodies against a specific epitope.

The precise design of a gp120 monomer in the receptor-stabilized conformation was possible because of the availability of a crystal structure of gp120 bound to CD4 and the CD4i antibody-17b. We expect atomic-level structures of the trimeric spike to enable similar precise design of immunogens with optimal exposure of stabilized vulnerable epitopes of the HIV envelope.

### 39.4 Methods for Determining Atomic-Level Structure

Critical to the utilization of structural information in HIV-1 vaccine design is the capacity to generate the requisite atomic-level information of the relevant molecules: gp120- and gp41-derived glycoproteins, HIV-1-reactive antibodies, and various immunogens. Two structural techniques – X-ray crystallography and NMR spectroscopy – are capable of generating such information. In our work to date, we have utilized X-ray crystallography almost exclusively, although collaborative efforts to obtain solution NMR structures have progressed [30, 56].

Elucidation of the trimeric HIV-1 Env spike structure is an important, though difficult, step towards rational design of an HIV-1 immunogen. Production of adequate amounts of highly purified protein is a necessary step towards obtaining crystals suitable for atomic-level analysis. We have assessed a number of expression systems for expressing HIV-1 Env glycoproteins suitable for structural analysis (Table 39.1), and currently employ a mammalian cell-based transient transfection system for protein production (Fig. 39.6). However, high-resolution structural analysis has been hindered by difficulties in producing adequate amounts of stable HIV-1 recombinant envelope trimers. Interestingly, poor viral spike expression has also been observed for other members of the class I family of viral fusion proteins, including HIV-2/SIV, coronaviruses, and retroviruses. To address this problem, we have used glycosidase processing inhibitors, multiple transfection reagents and several target cell types to optimize trimer expression, glycan processing, and cost efficiency (Fig. 39.6). At the moment, we are testing the effects of different signal sequences and trimerization domains on viral spike expression.

Even with the current state of the art technology for data collection, only crystals with an ordered internal structure can be used for generating atomic level information. Obtaining suitable crystals for the highly glycosylated, conformationally flexible gp120, is in the best cases demanding and in the worst cases not possible. We have pioneered a number of cutting-edge crystallization strategies to increase our chances of success (Fig. 39.7). These include variational protein crystallization [26], in which a number of closely related proteins are assessed for crystallization. Thus, for example, to obtain a complex of a CD4-induced antibody with gp120, 5 CD4-induced antibodies were combined with 3 different Clade B gp120s to make 15 different complexes. Each of the 15 complexes was then assessed for crystallization, enhancing the overall probability of obtaining crystals suitable for structural analysis [29]. We have also utilized efficient high-throughput methods for screening crystallization conditions, employing factorization [57], sparse matrices [58, 59], and precipitant-focused methods of iterative optimization; and we have also embraced the latest automated structural genomic advances [60] such as crystallization robots [61, 62] and synchrotron-based means of data collection [63, 64].

As protein flexibility may be a critical factor in elicitation, we have used hydrogen-deuterium exchange techniques [65] to quantify the flexibility of gp120 [66]. Information regarding conformational dynamics of a protein in solution can be probed using heteronuclear NMR spin relaxation spectroscopy [67–72]. These experiments require either uniformly or selectively labeled  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and/or  $^2\text{H}$  proteins, and we are currently optimizing protein expression in eukaryotic systems to obtain sufficient quantities of isotopically enriched HIV-1 envelope glycoprotein domains.

### 39.5 Ebola Virus

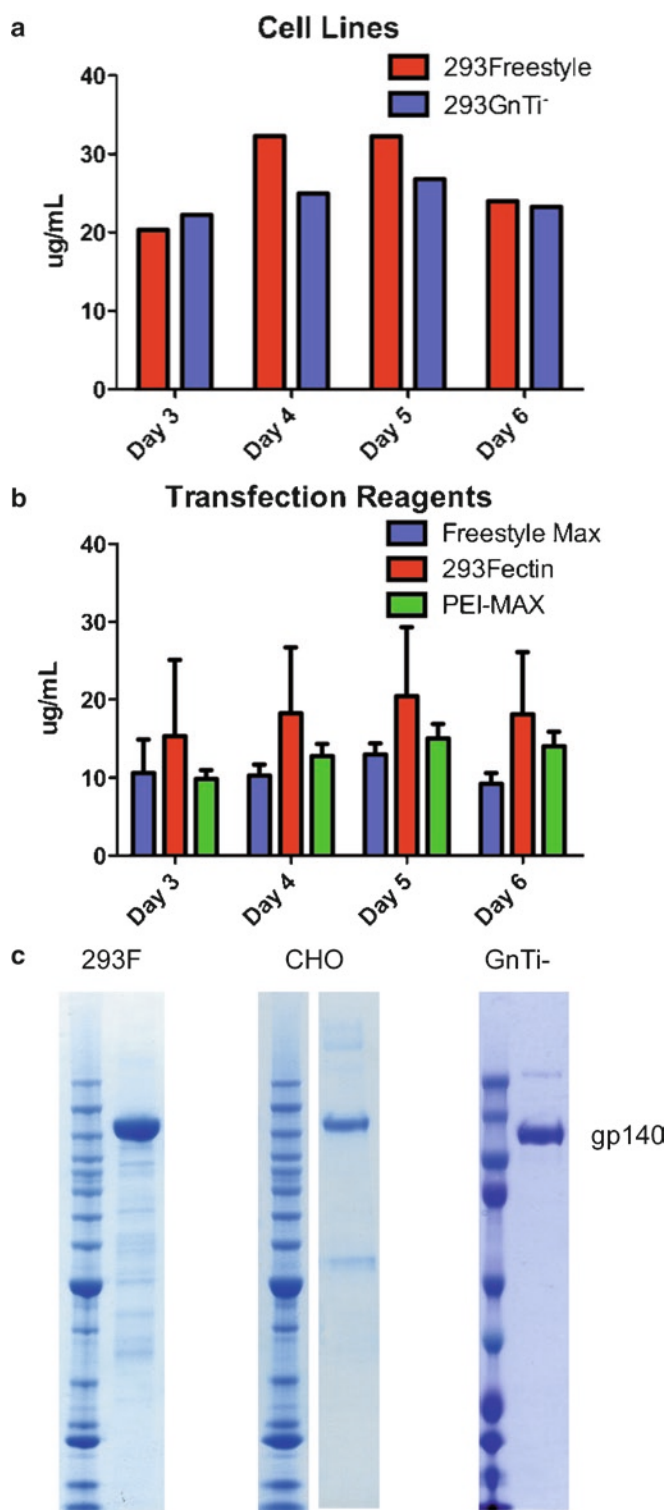
Because HIV-1 has numerous means of immune evasion, a vaccine will likely have to overcome extraordinary hurdles to be effective. To test the feasibility of clearing these hurdles, we have been attempting to apply structural biology to other viruses.

**Table 39.1** Production of HIV-1 HXBc2 gp120 core

Expression System	Time required to produce secreted protein starting from DNA construct	Typical yield of core gp120 after affinity purification
Drosophila – S2 cells	50–70 days*	5–10 mg/l
Adenovirus	40–50 days*	40–50 mg/l
Transient transfection – HEK293 Freestyle cells	3–6 days	10–20 mg/l

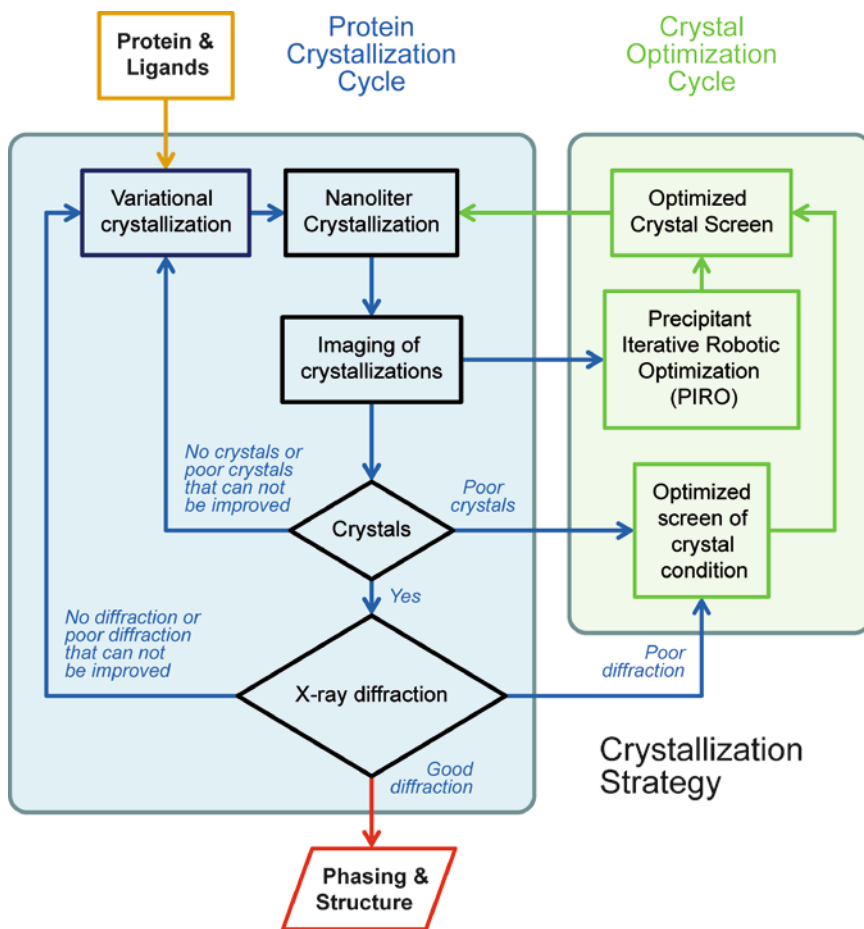
\*Includes time required to select a stable line and/or virus.





**Fig. 39.6** Effects of cell line and transfection reagents on HIV-1 gp140 production and purity. **(a)** Cell line. Transient transfections with 293Freestyle cells (Invitrogen) and 293Fectin (Invitrogen) had ~10% higher CAAN-gp140 yield compared to the 293GnTi-cells (a glycosylation-deficient suspension 293 cell line; HG Khorana) (87). **(b)** Transfection reagent. Transient transfections with 293GnTi-cells and 293Fectin had ~20% higher CAAN-gp140 yield as compared to transfections with Freestyle MAX (Invitrogen) and PEI-MAX (Polysciences) in the same cells. From these results, we conclude that

293GnTi-cells may be the most cost-effective option since glycosidase inhibitors are unnecessary and there is not a significant drop in yield. **(c)** Effect on the purity. A construct encoding a clade A gp140 (fibrin trimerization domain appended C-terminal to the gp41 ectodomain) was tested in transient transfections on 293Freestyle cells (293F), Chinese Hamster Ovary cells (CHO) and 293GnTi- cells (GnTi-). Transient transfections with 293GnTi-cells resulted in higher purity of the desired protein after a Ni<sup>2+</sup>-NTA column as compared to the 293Freestyle and CHO (Invitrogen) cells



**Fig. 39.7** Crystallization strategy. In the Protein Crystallization Cycle, the protein (P1) is complexed with ligands (L1). Variational crystallization, P1 is combined with different ligands (L1, L2, L3...) and/or different variations of the protein (P11, P12, P13...) are combined with ligands. Nanoliter crystallizations with droplets of total volume 200 nanoliters are set up with robots. For each complex, almost 600 different conditions are screened at the start. Data is imaged by robotic imaging systems and scored by the user. Suitable

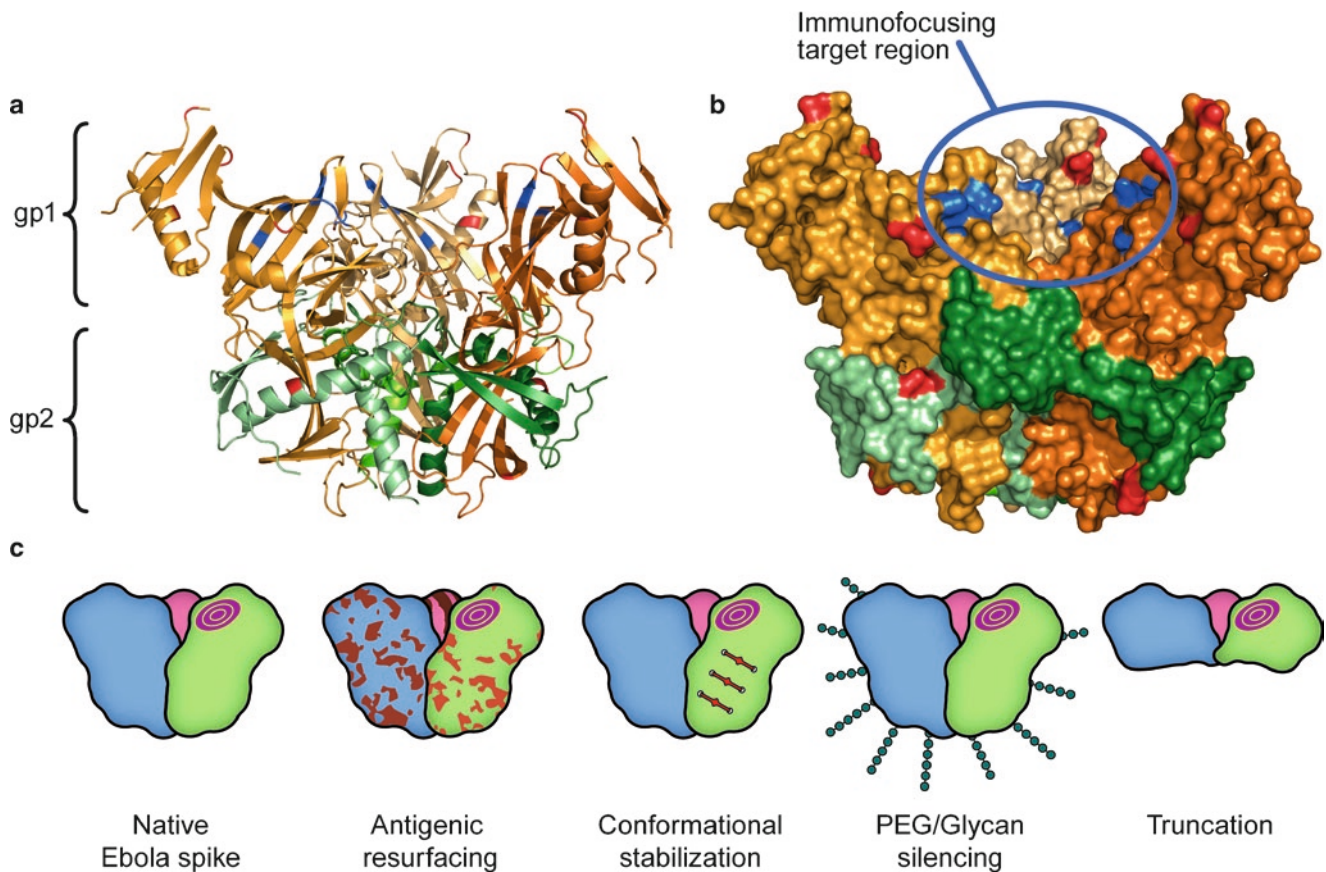
crystals are checked for x-ray diffraction. Good diffraction data is solved for the crystal structure of the protein or protein complex. The Crystal Optimization Cycle is used for optimizing the initial crystallization screening conditions, and for optimizing conditions which result in poor crystals. Imaging data is scored manually, which is used to generate a new crystallization matrix, referred to as an optimized screen. This is based on the optimizing of only the precipitant within each condition of the screen

Ebola virus is an enveloped, negative-strand RNA virus, belonging to the *Filoviridae* family. Ebola virus infection causes severe haemorrhagic fever, and depending on the strain, results in up to 90% mortality. A single envelope glycoprotein on the virion surface is responsible for attachment and entry into permissive cells, presumably via receptor-mediated endocytosis. Following endocytosis, the glycoprotein is cleaved by cathepsin B and/or cathepsin L in the acidic endosome environment and potentially triggers membrane fusion, which subsequently allows for the entry of the Ebola nucleocapsid into the cell cytoplasm [73].

The Ebola glycoprotein is generated as a precursor GP0 protein, which is cleaved at a furin-like site to yield GP1 and GP2. These proteins are linked by a disulphide bond, and a trimer complex of this heterodimer forms the viral spike. Recently, the crystal structure of prefusion Ebola virus strain

Zaire was determined in complex with the human neutralizing antibody KZ52 [3]. The structure revealed that GP1 possesses an open chalice-like shape, while GP2 forms a belt around the base to create intimate GP1-GP2 and GP2-GP2 contacts (Fig. 39.8a). While a protective vaccine against Ebola will likely require the elicitation of an appropriate cellular immune response, neutralizing antibody responses against the Ebola viral spike may also play a significant protective role.

The recent structure determination of the Ebola Zaire glycoprotein provides a blueprint to design immunogens that are targeted to biologically relevant regions on the structure (Fig. 39.8c). One of the methods that can be employed to immunofocus the response is silencing regions that are not biologically relevant or known to elicit an unfavorable immune response. For example, the human neutralizing antibody, KZ52, binds to GP1:GP2 residues in the base region



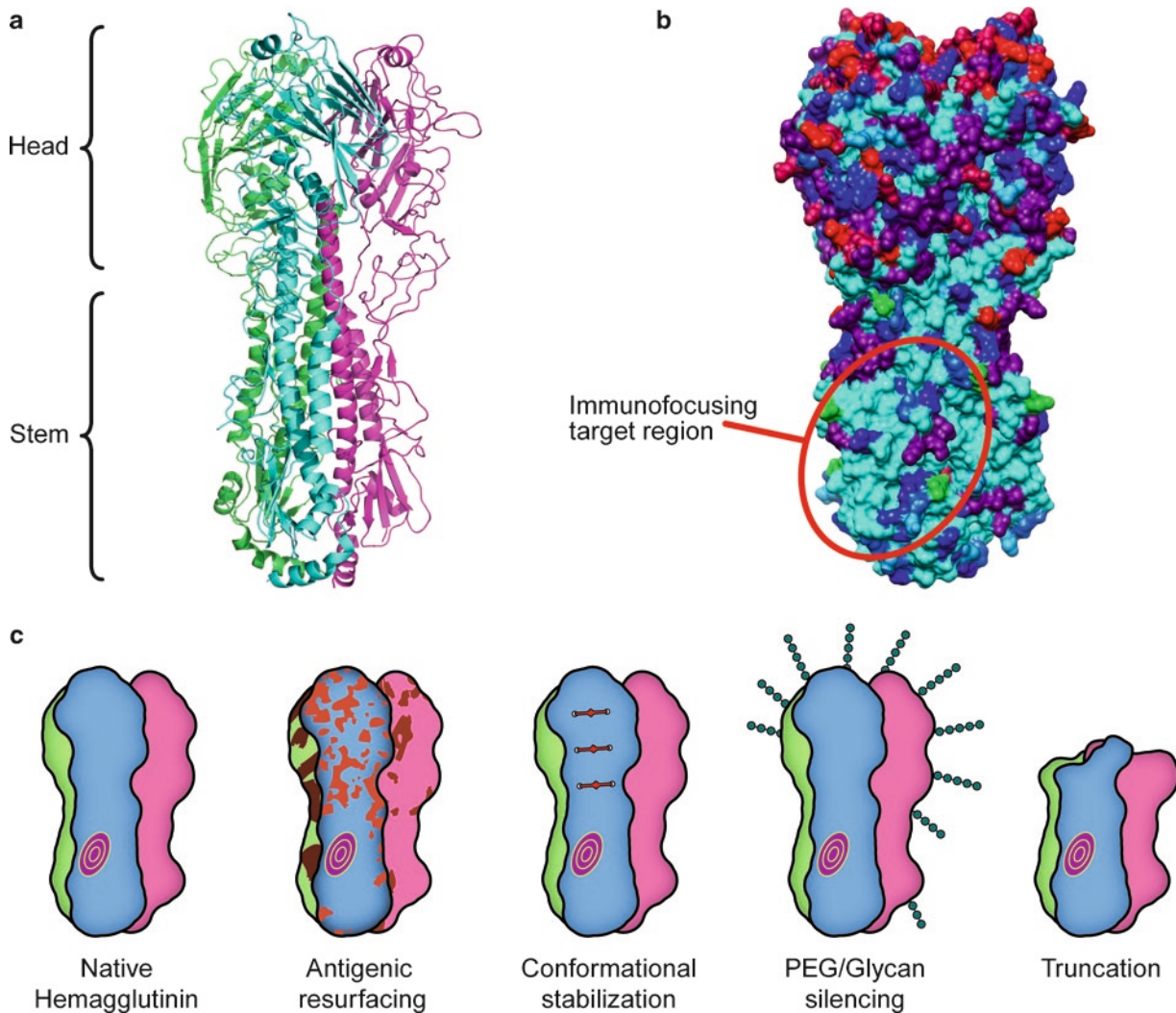
**Fig. 39.8** Ebola viral spike: trimeric ectodomain structure and immunofocusing methods. **(a)** Trimeric structure. The crystal structure (3) of the ectodomain of the Ebola Zaire prefusion viral spike is depicted in C $\alpha$ -backbone representation. It adopts an open chalice-like shape of GP1 (each monomer is a different shade of orange), held together by a belt of GP2 (monomers shown in various shades of green). The receptor-binding domain is localized between residues 54 and 201 of GP1 (88), of which 6 residues of known critical importance for virus entry have been mapped onto the structure (shown in blue). These residues are located in the head region, which itself is surrounded by a glycan cap containing

several N-linked glycan sites (shown in red). The mucin-like domain of GP1, which was not included in the protein crystallized, would be modeled to surround the glycan cap and further extensively glycosylate the protein with both N- and O-linked sugars. Coordinates from (3); PDB ID 3csy. **(b)** Target surface. The surface of the Ebola viral spike ectodomain is shown from the same coloring and orientation as in (a), with the putative receptor-binding region highlighted. **(c)** Immunofocusing strategies. The target region for immunofocusing methods is designated by a “bull’s eye” in the left-most image, and schematics for four immunofocusing strategies are depicted

of the trimer [3]. We can effectively focus the immunogen to elicit antibodies to the conserved trimer core through the addition of glycans to specific residues to which KZ52 binds. Another approach is to remove the highly glycosylated mucin-like region, which may play a role in providing the ultimate virus immune evasion strategy. Removal of this region also focuses the immune response to the exposed highly conserved receptor binding core. Our current immunogen approach, to target the highly conserved receptor-binding domain, may also result in generating protection against many, if not all, Ebola virus strains. Knowledge of the crystal structure of the prefusion Ebola glycoprotein also allows us to create recombinant proteins that mimic the processed glycoprotein that may represent a conserved but otherwise inaccessible form of the protein that may be sensitive to neutralization.

## 39.6 Influenza Virus

Influenza virus results in 3–5 million cases of severe illness per year causing up to 500,000 deaths worldwide (WHO EB111/10) with the most severe cases occurring in young children and the elderly. In addition to humans, influenza also infects numerous species of mammals and birds, although wild waterfowl are thought to be the primary reservoir [74]. Influenza is a spherically-shaped single-strand negative sense RNA virus belonging to the *Orthomyxoviridae* family. The outer viral surface comprises three membrane-anchored proteins: hemagglutinin (HA), neuraminidase (NA) and M2. HA is the most abundant and immunogenic of the three. To date, all neutralizing monoclonal antibodies to influenza target HA; no neutralizing antibodies against NA or M2 have been reported.



**Fig. 39.9** Influenza hemagglutinin: trimeric ectodomain structure and immunofocusing methods. (a) Ribbon diagram of an H5N1 hemagglutinin trimer. The HA1-HA2 subunits from each hemagglutinin protomer are colored blue, green and magenta respectively. Coordinates from (75); PDB ID 2fk0. (b) Sequence variation between H1, H2 and H5 hemagglutinins mapped onto the molecular surface. Cyan depicts completely conserved and red highly variable amino acids. Color gradations

between cyan and red represent increasing degrees of amino acid variability. Green represents Asn residues of conserved N-linked glycosylation sequons. A suggested region for immunofocusing efforts is highlighted with a red circle. (c) Immunofocusing strategies. The target region for immunofocusing methods is designated by a “bull’s eye” in the left-most image, and schematics for four immunofocusing strategies are depicted

Crystal structures of H1, H3, H5, H7, H9 and B HAs have each revealed a 13.5 nm cylinder-shaped homotrimeric glycoprotein consisting of a mostly  $\beta$ -sheet globular head sitting atop a predominantly helical stem region (Fig. 39.9a) [4, 75, 76]. HA is a class I fusion protein that is cleaved during expression to generate HA1 and an HA2 with an N-terminal fusion peptide. The head region of HA recognizes terminal sialic acid (i.e. the receptor) on the surface of the host cell to facilitate viral internalization. Upon acidification below pH 6.0 in the endosome, HA undergoes a dramatic conformational change that promotes insertion of the fusion peptide into the host membrane and brings the viral and host

transmembrane regions of HA into close enough proximity to induce membrane fusion.

Influenza comprises three main viruses: A, B and C, although only A and B cause disease in humans. Type-B has no subtypes. However, influenza A is highly variable and currently consists of 16 known immunological subtypes of HA and 9 subtypes of NA combined in a multitude of reassorted strains. To date, most human infections have resulted from H1N1, H1N2, H2N2, and H3N2 strains, although more recently 390 cases of highly pathogenic H5N1 infections and a handful of H7 and H9 infections of humans have been reported [77]. In total, over 6,000 different strains of influenza

are now known to infect humans [78]. Influenza is highly immunogenic; vaccines created from inactivated or live-attenuated virus are generally effective without the need for a booster shot or the use of adjuvants. Unfortunately, selective pressure from the immune system of hosts combined with an error prone viral RNA polymerase promotes rapid mutation of the influenza viral surface proteins (referred to as antigenic drift) as well as genetic reassortments. As a result, new trivalent vaccines need to be generated each year using the strains of H1N1, H3N2 and B-type influenza that are predicted to be in circulation six months later during the upcoming flu season [79]. Not only does this require a significant outpouring of time and resources every year, but the predictions regarding upcoming influenza strains are not always on target. Furthermore, a large unpredicted human outbreak from a pathogenic avian subtype such as H5N1 would not be covered by current vaccines.

We are therefore focusing our efforts to design a vaccine that can direct the immune response, particularly antibodies, against more conserved regions of the influenza virus that are not likely to mutate quickly. In 1993, Okuno and colleagues reported the generation of a broadly neutralizing murine antibody that could neutralize H1, H2, H5 and some H6 HAs [80]. More recently Kashyap and colleagues reported on broadly neutralizing human monoclonal antibodies that neutralize divergent H1 and H5 strains [81]. These reports suggest that conserved regions of HA can elicit broadly neutralizing antibodies. When sequence conservation is mapped onto the three-dimensional structure of HA, it is evident that the stem region consisting mostly of HA2 is much more conserved than the HA1-head region (Fig. 9b). Since most neutralizing HA epitopes are located on the highly variable head region, we are presently exploring methods to immunofocus antibodies to the conserved HA-stem. The methods we are employing include PEG/glycan silencing, antigenic resurfacing and protein truncation of immunodominant regions as well as removal of native glycosylation and conformational stabilization (Fig. 9c). Our ultimate goal is to design an influenza immunogen that will elicit antibodies that broadly neutralize not only against multiple strains but even against multiple subtypes of influenza over time spans of many years.

### 39.7 Respiratory Syncytial Virus

Human respiratory syncytial virus (RSV) is a highly contagious pathogen, infecting almost the entire US population by the second year of life [82]. It is the leading cause of bronchiolitis and pneumonia in children less than 12 months old, and it is responsible for ~100,000 hospitalizations each year in the U.S. for children in this age group [82]. RSV can lead to

severe disease and death for premature infants and young children with congenital heart disease or compromised immune systems. The elderly are also susceptible to RSV infection, which results in more than 9,000 deaths per year in the U.S. [83].

Currently, there is no approved vaccine for RSV. A formalin-inactivated virus preparation was found to elicit an immune response that enhanced, rather than reduced, disease symptoms [84]. Infections in the highest risk groups can be prevented with monthly injections of neutralizing monoclonal antibodies during the RSV season, but this is an expensive treatment that is not readily available to all people in the world. As RSV shares some similarities with HIV-1 at the molecular level, it is an appealing target for structure-based vaccine design.

The F protein of RSV is related to the gp120/gp41 viral spike protein of HIV-1, and mediates fusion of the viral and cellular membranes. It is also the target of several neutralizing antibodies, including those currently being used for disease prevention. Atomic-level structures of trimeric F proteins from related viruses in both pre-fusion and post-fusion conformations exist [5, 85, 86], allowing for a molecular understanding of interactions within the viral spike that is currently not available for HIV-1. We hope to apply many of the methods and techniques originally developed for creating HIV-1 immunogens to the design of an effective RSV F protein-based vaccine. These include the attachment of heterologous trimerization domains to the C-terminus of the RSV viral spike, to stabilize it in its prefusion conformation. Other efforts include the use of epitope-scaffold technologies to focus the immune response onto appropriate areas. Definition of beneficial immune responses and their mechanisms of protection against RSV may be necessary to prevent deleterious immunization effects.

### 39.8 Summary

The intramural program in the Vaccine Research Center of the National Institutes of Allergy and Infectious Diseases at NIH provides an excellent setting in which to combine basic discovery with translational efforts. Whether the confluence of structural information with protein design and immune analysis is sufficient to elicit broadly neutralizing antibodies will depend in part on our ability to optimize the immune response and also on the rationale of conformational mimicry, epitope accessibility, neutralization breadth, and target specificity. Our investigations have already provided insight into the parameters governing antibody elicitation and neutralization. Whether such basic science discoveries will provide benefits to public health depends on our ability to translate them from the laboratory to the clinic, a critical step facilitated

by close interactions with other laboratories at the center. True success with HIV-1 will ultimately depend on whether we succeed in creating immunogens that can elicit broadly neutralizing responses which can then be translated into vaccine regimens capable of substantially reducing the incidence of HIV-1 infection in humans. Other future and past pandemic pathogens such as Ebola virus and influenza A virus may also be surmounted with structure-assisted efforts.

## References

- Zhou, T., Xu, L., Dey, B., Hessel, A.J., Van Ryk, D., Xiang, S.H., Yang, X., Zhang, M.Y., Zwick, M.B., Arthos, J., Burton, D.R., Dimitrov, D.S., Sodroski, J., Wyatt, R., Nabel, G.J. and Kwong, P.D. (2007). Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* 445:732–737.
- Li, Y., Migueles, S.A., Welcher, B., Svehla, K., Phogat, A., Louder, M.K., Wu, X., Shaw, G.M., Connors, M., Wyatt, R.T. and Mascola, J.R. (2007). Broad HIV-1 neutralization mediated by CD4-binding site antibodies. *Nat Med* 13:1032–1034.
- Lee, J.E., Fusco, M.L., Hessel, A.J., Oswald, W.B., Burton, D.R. and Saphire, E.O. (2008). Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454:177–182.
- Wilson, I.A., Skehel, J.J. and Wiley, D.C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289:366–373.
- Yin, H.-S., Wen, X., Paterson, R.G., Lamb, R.A. and Jardetzky, T.S. (2006). Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* 439:38–44.
- Harrison, S.C. (2005). Mechanism of membrane fusion by viral envelope proteins. *Adv Virus Res* 64:231–261.
- Colman, P.M. and Lawrence, M.C. (2003). The structural biology of type I viral membrane fusion. *Nat Rev Mol Cell Biol* 4:309–319.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S. and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* 288:1789–1796.
- (2008). 2008 Report on the global AIDS Epidemic, in Joint United Nations Programme on HIV/AIDS.
- Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A. and Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* 393:705–711.
- Wyatt, R. and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280:1884–1888.
- Roux, K.H. and Taylor, K.A. (2007). AIDS virus envelope spike structure. *Curr Opin Struct Biol* 17:244–252.
- Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G. and Subramaniam, S. (2008). Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109–113.
- Kwong, P.D. (2005). Human immunodeficiency virus: refolding the envelope. *Nature* 433:815–816.
- Chen, B., Vogan, E.M., Gong, H., Skehel, J.J., Wiley, D.C. and Harrison, S.C. (2005). Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* 433:834–841.
- Feng, Y., Broder, C.C., Kennedy, P.E. and Berger, E.A. (1996). HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272:872–877.
- Moore, J.P. (1997). Coreceptors: implications for HIV pathogenesis and therapy. *Science* 276:51–52.
- Berger, E.A., Murphy, P.M. and Farber, J.M. (1999). Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* 17:657–700.
- Caffrey, M., Cai, M., Kaufman, J., Stahl, S.J., Wingfield, P.T., Covell, D.G., Gronenborn, A.M. and Clore, G.M. (1998). Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J* 17:4572–4584.
- Lu, M., Blacklow, S.C. and Kim, P.S. (1995). A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol* 2:1075–1082.
- Chan, D.C., Fass, D., Berger, J.M. and Kim, P.S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89:263–273.
- Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J. and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387:426–430.
- Tan, K., Liu, J., Wang, J., Shen, S. and Lu, M. (1997). Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci U S A* 94:12303–12308.
- Yang, X., Lee, J., Mahony, E.M., Kwong, P.D., Wyatt, R. and Sodroski, J. (2002). Highly stable trimers formed by human immunodeficiency virus type 1 envelope glycoproteins fused with the trimeric motif of T4 bacteriophage fibritin. *J Virol* 76:4634–4642.
- Kim, S., Pang, H.B. and Kay, M.S. (2008). Peptide mimic of the HIV envelope gp120-gp41 interface. *J Mol Biol* 376:786–797.
- Kwong, P.D., Wyatt, R., Desjardins, E., Robinson, J., Culp, J.S., Hellmig, B.D., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (1999). Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1). *J Biol Chem* 274:4115–4123.
- Kwong, P.D., Wyatt, R., Majeed, S., Robinson, J., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (2000). Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure* 8:1329–1339.
- Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393:648–659.
- Huang, C.C., Tang, M., Zhang, M.Y., Majeed, S., Montabana, E., Stanfield, R.L., Dimitrov, D.S., Korber, B., Sodroski, J., Wilson, I.A., Wyatt, R. and Kwong, P.D. (2005). Structure of a V3-containing HIV-1 gp120 core. *Science* 310:1025–1028.
- Huang, C.C., Lam, S.N., Acharya, P., Tang, M., Xiang, S.H., Hussan, S.S., Stanfield, R.L., Robinson, J., Sodroski, J., Wilson, I.A., Wyatt, R., Bewley, C.A. and Kwong, P.D. (2007). Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science* 317:1930–1934.
- Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., Komarova, N.L., Nowak, M.A., Hahn, B.H., Kwong, P.D. and Shaw, G.M. (2003). Antibody neutralization and escape by HIV-1. *Nature* 422:307–312.
- Wyatt, R., Moore, J., Accola, M., Desjardin, E., Robinson, J. and Sodroski, J. (1995). Involvement of the V1/V2 variable loop structure in the exposure of human immunodeficiency virus type 1 gp120 epitopes induced by receptor binding. *J Virol* 69:5723–5733.
- Kwong, P.D., Doyle, M.L., Casper, D.J., Cicala, C., Leavitt, S.A., Majeed, S., Steenbeke, T.D., Venturi, M., Chaiken, I., Fung, M., Katinger, H., Parren, P.W., Robinson, J., Van Ryk, D., Wang, L., Burton, D.R., Freire, E., Wyatt, R., Sodroski, J., Hendrickson, W.A. and Arthos, J. (2002). HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 420:678–682.
- Decker, J.M., Bibollet-Ruche, F., Wei, X., Wang, S., Levy, D.N., Wang, W., Delaporte, E., Peeters, M., Derdeyn, C.A., Allen, S., Hunter, E., Saag, M.S., Hoxie, J.A., Hahn, B.H., Kwong, P.D., Robinson, J.E. and Shaw, G.M. (2005). Antigenic conservation and immunogenicity of the HIV coreceptor binding site. *J Exp Med* 201:1407–1419.

35. Salzwedel, K., Smith, E.D., Dey, B. and Berger, E.A. (2000). Sequential CD4-coreceptor interactions in human immunodeficiency virus type 1 Env function: soluble CD4 activates Env for coreceptor-dependent fusion and reveals blocking activities of antibodies against cryptic conserved epitopes on gp120. *J Virol* 74:326–333.
36. Labrijn, A.F., Poignard, P., Raja, A., Zwick, M.B., Delgado, K., Franti, M., Binley, J., Vivona, V., Grundner, C., Huang, C.C., Venturi, M., Petropoulos, C.J., Wrin, T., Dimitrov, D.S., Robinson, J., Kwong, P.D., Wyatt, R.T., Sodroski, J. and Burton, D.R. (2003). Access of antibody molecules to the conserved coreceptor binding site on glycoprotein gp120 is sterically restricted on primary human immunodeficiency virus type 1. *J Virol* 77:10557–10565.
37. Burton, D.R., Pyati, J., Koduri, R., Sharp, S.J., Thornton, G.B., Parren, P.W., Sawyer, L.S., Hendry, R.M., Dunlop, N., Nara, P.L. and et al. (1994). Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science* 266:1024–1027.
38. Muster, T., Guinea, R., Trkola, A., Purtscher, M., Klima, A., Steindl, F., Palese, P. and Katinger, H. (1994). Cross-neutralizing activity against divergent human immunodeficiency virus type 1 isolates induced by the gp41 sequence ELDKWAS. *J Virol* 68:4031–4034.
39. Muster, T., Steindl, F., Purtscher, M., Trkola, A., Klima, A., Himmler, G., Ruker, F. and Katinger, H. (1993). A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* 67:6642–6647.
40. Trkola, A., Purtscher, M., Muster, T., Ballaun, C., Buchacher, A., Sullivan, N., Srinivasan, K., Sodroski, J., Moore, J.P. and Katinger, H. (1996). Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J Virol* 70:1100–1108.
41. Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J.R., Wyatt, R. and Kwong, P.D. (2004). Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol* 78:10724–10737.
42. Cavacini, L.A., Samore, M.H., Gambertoglio, J., Jackson, B., Duval, M., Wisniewski, A., Hammer, S., Koziel, C., Trapnell, C. and Posner, M.R. (1998). Phase I study of a human monoclonal antibody directed against the CD4-binding site of HIV type 1 glycoprotein 120. *AIDS Res Hum Retroviruses* 14:545–550.
43. Lin, P.F., Blair, W., Wang, T., Spicer, T., Guo, Q., Zhou, N., Gong, Y.F., Wang, H.G., Rose, R., Yamanaka, G., Robinson, B., Li, C.B., Fridell, R., Deminie, C., Demers, G., Yang, Z., Zadajura, L., Meanwell, N. and Colonna, R. (2003). A small molecule HIV-1 inhibitor that targets the HIV-1 envelope and inhibits CD4 receptor binding. *Proc Natl Acad Sci U S A* 100:11013–11018.
44. Zhao, Q., Ma, L., Jiang, S., Lu, H., Liu, S., He, Y., Strick, N., Neamati, N. and Debnath, A.K. (2005). Identification of N-phenyl-N'-(2,2,6,6-tetramethyl-piperidin-4-yl)-oxalamides as a new class of HIV-1 entry inhibitors that prevent gp120 binding to CD4. *Virology* 339:213–225.
45. Madani, N., Schon, A., Princiotta, A.M., Lalonde, J.M., Courter, J.R., Soeta, T., Ng, D., Wang, L., Brower, E.T., Xiang, S.H., Do Kwon, Y., Huang, C.C., Wyatt, R., Kwong, P.D., Freire, E., Smith, A.B., 3rd and Sodroski, J. (2008). Small-molecule CD4 mimics interact with a highly conserved pocket on HIV-1 gp120. *Structure* 16:1689–1701.
46. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., Roumestand, C., Yang, Y.S., Ylisastigui, L., Benjouad, A. and Gluckman, J.C. (1999). Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein. *Proc Natl Acad Sci USA* 96:13091–13096.
47. Martin, L., Stricher, F., Misse, D., Sironi, F., Pugniere, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., Menez, A., Lusso, P., Veas, F. and Vita, C. (2003). Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. *Nat Biotechnol* 21:71–76.
48. Stricher, F., Huang, C.C., Descours, A., Duquesnoy, S., Combes, O., Decker, J.M., Kwon, Y.D., Lusso, P., Shaw, G.M., Vita, C., Kwong, P.D. and Martin, L. (2008). Combinatorial optimization of a CD4-mimetic miniprotein and cocystal structures with HIV-1 gp120 envelope glycoprotein. *J Mol Biol* 382:510–524.
49. Huang, C.C., Stricher, F., Martin, L., Decker, J.M., Majeed, S., Barthe, P., Hendrickson, W.A., Robinson, J., Roumestand, C., Sodroski, J., Wyatt, R., Shaw, G.M., Vita, C. and Kwong, P.D. (2005). Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. *Structure* 13:755–768.
50. Pantophlet, R., Wilson, I.A. and Burton, D.R. (2004). Improved design of an antigen with enhanced specificity for the broadly HIV-neutralizing antibody b12. *Protein Eng Des Sel* 17:749–758.
51. Pantophlet, R., Wilson, I.A. and Burton, D.R. (2003). Hyperglycosylated mutants of human immunodeficiency virus (HIV) type 1 monomeric gp120 as novel antigens for HIV vaccine design. *J Virol* 77:5889–5901.
52. Roberts, M.J., Bentley, M.D. and Harris, J.M. (2002). Chemistry for peptide and protein PEGylation. *Adv Drug Deliver Rev* 54:459–476.
53. O'Riordan, C.R., Lachapelle, A., Delgado, C., Parkes, V., Wadsworth, S.C., Smith, A.E. and Francis, G.E. (1999). PEGylation of adenovirus with retention of infectivity and protection from neutralizing antibody in vitro and in vivo. *Hum Gene Ther* 10:1349–1358.
54. Ogris, M., Brunner, S., Schuller, S., Kircheis, R. and Wagner, E. (1999). PEGylated DNA/transferrin-PEI complexes: reduced interaction with blood components, extended circulation in blood and potential for systemic gene delivery. *Gene Ther* 6:595–605.
55. Dey, B., Svehla, K., Xu, L., Wycuff, D., Zhou, T., Voss, G., Phogat, A., Chakrabarti, B.K., Li, Y., Shaw, G.M., Kwong, P.D., Nabel, G.J., Mascola, J. and Wyatt, R.T. (2009). Structure-based stabilization of HIV-1 gp120 enhances humoral immune responses to the induced co-receptor binding site. Submitted.
56. Lam, S.N., Acharya, P., Wyatt, R., Kwong, P.D. and Bewley, C.A. (2008). Tyrosine-sulfate isosteres of CCR5 N-terminus as tools for studying HIV-1 entry. *Bioorg Med Chem*:In press.
57. Carter, C.W., Jr. and Carter, C.W. (1979). Protein crystallization using incomplete factorial experiments. *J Biol Chem* 254:12219–12223.
58. Jancarik, J. and Kim, S.H. (1991). Sparse-Matrix Sampling - a Screening Method for Crystallization of Proteins. *J Appl Crystallogr* 24:409–411.
59. Majeed, S., Ofek, G., Belachew, A., Huang, C.C., Zhou, T. and Kwong, P.D. (2003). Enhancing protein crystallization through precipitant synergy. *Structure* 11:1061–1070.
60. Stevens, R.C. and Wilson, I.A. (2001). Tech.Sight. Industrializing structural biology. *Science* 293:519–520.
61. Hosfield, D., Palan, J., Hilgers, M., Scheibe, D., McRee, D.E. and Stevens, R.C. (2003). A fully integrated protein crystallization platform for small-molecule drug discovery. *J Struct Biol* 142:207–217.
62. Morris, D.W., Kim, C.Y. and McPherson, A. (1989). Automation of protein crystallization trials: use of a robot to deliver reagents to a novel multi-chamber vapor diffusion plate. *Biotechniques* 7:522–527.
63. Hendrickson, W.A. (2000). Synchrotron crystallography. *Trends in Biochemical Sciences* 25:637–643.
64. Sakabe, N. (1991). X-Ray-Diffraction Data-Collection System for Modern Protein Crystallography with a Weissenberg Camera and an Imaging Plate Using Synchrotron Radiation. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* 303:448–463.
65. Zhang, Z.Q. and Smith, D.L. (1993). Determination of Amide Hydrogen-Exchange by Mass-Spectrometry - a New Tool for Protein-Structure Elucidation. *Protein Sci* 2:522–531.

66. Smith, D.L., Deng, Y.Z. and Zhang, Z.Q. (1997). Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J Mass Spectrom* 32:135–146.
67. Cavanagh, J., Fairbrother, W.J., Palmer 3rd, A.G., Rance, M. and Skelton, N.J. (2007). *Protein NMR Spectroscopy: Principles and Practice*. Elsevier, Oxford.
68. Igumenova, T.I., Frederick, K.K. and Wand, A.J. (2006). Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chemical Reviews* 106:1672–1699.
69. Jarymowycz, V.A. and Stone, M.J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chemical Reviews* 106:1624–1671.
70. Palmer, A.G. and Massi, F. (2006). Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. *Chemical Reviews* 106:1700–1719.
71. Siggers, K., Soto, C. and Palmer, 3rd A.G. (2007). Conformational dynamics in loop swap mutants of homologous fibronectin type III domains. *Biophys J* 93:2447–2456.
72. Tolman, J.R. and Ruan, K. (2006). NMR Residual Dipolar Couplings as Probes of Biomolecular Dynamics. *Chemical Reviews* 106:1720–1736.
73. Schornberg, K., Matsuyama, S., Kabsch, K., Delos, S., Bouton, A. and White, J. (2006). Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol* 80:4174–4178.
74. Kovacova, A., Ruttkay-Nedecky, G., Haverlik, I.K. and Janecek, S. (2002). Sequence similarities and evolutionary relationships of influenza virus A hemagglutinins. *Virus Genes* 24:57–63.
75. Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C. and Wilson, I.A. (2006). Structure and Receptor Specificity of the Hemagglutinin from an H5N1 Influenza Virus. *Science* 312:404–410.
76. Wang, Q., Tian, X., Chen, X. and Ma, J. (2007). Structural basis for receptor specificity of influenza B virus hemagglutinin. *Proc Natl Acad Sci USA* 104:16874–16879.
77. WHO (2008). [http://www.who.int/csr/disease/avian\\_influenza/country/cases\\_table\\_2008\\_12\\_12/en/index.html](http://www.who.int/csr/disease/avian_influenza/country/cases_table_2008_12_12/en/index.html).
78. Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., Ostell, J. and Lipman, D. (2008). The Influenza Virus Resource at the National Center for Biotechnology Information. *J. Virol.* 82:596–601.
79. Nichol, K.L. and Treanor, J.J. (2006). Vaccines for Seasonal and Pandemic Influenza. *J Infect Dis* 194:S111–S118.
80. Okuno, Y., Isegawa, Y., Sasao, F. and Ueda, S. (1993). A common neutralizing epitope conserved between the hemagglutinins of influenza A virus H1 and H2 strains. *J. Virol.* 67:2552–2558.
81. Kashyap, A.K., Steel, J., Oner, A.F., Dillon, M.A., Swale, R.E., Wall, K.M., Perry, K.J., Faynboym, A., Ilhan, M., Horowitz, M., Horowitz, L., Palese, P., Bhatt, R.R. and Lerner, R.A. (2008). Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies. *Proc Natl Acad Sci USA* 105:5986–5991.
82. CDC (2008). <http://www.cdc.gov/rsv/index.html>.
83. Thompson, W.W., Shay, D.K., Weintraub, E., Brammer, L., Cox, N., Anderson, L.J. and Fukuda, K. (2003). Mortality Associated With Influenza and Respiratory Syncytial Virus in the United States. *JAMA* 289:179–186.
84. Kim, H.W., Canchola, J.G., Brandt, C.D., Pyles, G., Chanock, R.M., Jensen, K. and Parrott, R.H. (1969). Respiratory Syncytial Virus Disease in Infants Despite Prior Administration of Antigenic Inactivated Vaccine. *Am. J. Epidemiol.* 89:422–434.
85. Chen, L., Gorman, J.J., McKimm-Breschkin, J., Lawrence, L.J., Tulloch, P.A., Smith, B.J., Colman, P.M. and Lawrence, M.C. (2001). The Structure of the Fusion Glycoprotein of Newcastle Disease Virus Suggests a Novel Paradigm for the Molecular Mechanism of Membrane Fusion. *Structure* 9:255–266.
86. Yin, H.-S., Paterson, R.G., Wen, X., Lamb, R.A. and Jardetzky, T.S. (2005). Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proc Natl Acad Sci USA* 102:9288–9293.
87. Reeves, P.J., Callewaert, N., Contreras, R. and Khorana, H.G. (2002). Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl. Acad. Sci U.S.A.* 99:13419–13424.
88. Kuhn, J.H., Radoshitzky, S.R., Guth, A.C., Warfield, K.L., Li, W., Vincent, M.J., Towner, J.S., Nichol, S.T., Bavari, S., Choe, H., Aman, M.J. and Farzan, M. (2006). Conserved Receptor-binding Domains of Lake Victoria Marburgvirus and Zaire Ebolavirus Bind a Common Receptor. *J. Biol. Chem.* 281:15951–15958.