INVESTIGATIONS IN THE DESIGN AND CHARACTERIZATION

OF HIV-1 NEUTRALIZING MOLECULES

Thesis by

Joshua Simon Klein

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 28, 2009)

This work is dedicated to the memory of my father,

RONALD ALLEN KLEIN

(1947 – 1988)

ACKNOWLEDGEMENTS

It is often said, "There is no 'I' in team." That observation is especially accurate when applied to the work presented in this thesis as it constituted an effort that far exceeded what could be accomplished with one brain and two hands. My graduate experience was blessed by the support of an outstandingly committed and tireless team of research technicians that included Priyanthi Gnanapragasm, Rachel Galimidi, Maria Suzuki, Kathryn Huey-Tubman, and the members of the Protein Expression Center, headed by Dr. Jost Vielmetter, which includes Chris Foglesong, Inder Nagiana, and Michael Anaya. Without their help, this work would have easily taken many more years. In addition to their technical assistance, I am deeply grateful for their support, encouragement, and perennial smiles. My thanks also go to Dr. Anthony West and the rest of my fellow laboratory members, both past and present, whose technical advice has been invaluable.

I am indebted to my advisor, Dr. Pamela J. Björkman, for inviting me to join her laboratory at the inception of this project, for her dedication to the cultivation of my potential as a scientist, and for giving me the freedom to pursue every idea that I felt was worthy of investigation, including the ones that may not have been.

Finally, I would like to express my heartfelt thanks to and admiration for the Bill and Melinda Gates Foundation, whose financial support made this work possible.

ABSTRACT

Human Immunodeficiency Virus (HIV) is a T-lymphotrophic retrovirus that is the causative agent of Acquired Immunodeficiency Syndrome and is estimated to currently infect approximately 40 million people worldwide. Life-extending therapies are credited for the precipitous drop in HIV-related mortality in developed countries, but their high costs prevent widespread distribution in developing countries. To date, all attempts to produce a vaccine capable of preventing or controlling an HIV infection have failed, but a comprehensive explanation for these failures has yet to emerge from the available data. In this thesis the first chapter provides an overview of the pandemic, the antigenic properties of gp120 and gp41, which are the two glycoproteins that comprise the outer envelope spike of the virus, and the broadly neutralizing antibodies that have been isolated against them. The second and third chapters discuss biophysical characterizations of these monoclonal antibodies and newly designed molecules derived from them. Based on a comparison of these data with pre-existing research, a novel hypothesis called the "island effect" was developed and is presented as a possible explanation for the consistent failure of the human immune system to respond to infection or vaccination with an effective humoral response. The final chapter summarizes ongoing investigations in the capacities of broadly neutralizing monoclonal antibodies to recruit antibody-dependent cellular cytotoxicity, a mechanism by which antibodies can trigger the lysis of HIV-infected cells by the innate immune system.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

x

CHAPTER 4:

Figure 1. Four cell lines examined for stable surface expression of gp160.................93

Figure 2. Summary of curves for the in vitro neutralization data and *in vitro* ADCC data...96

Figure 3. Summary of curves from the in vitro ADCC data for control tests of IgG1 2G12 monomer and dimer versus untransfected CHO cells 97

Figure 4. Model structure of IgG-CD16 complex 100

Figure 5. Conceptual model of the distance requirements for simultaneous binding of IgG1 with CD16 and the HIV envelope spike with the assumption that the 16-residue unstructured linker connecting the extracellular domains of CD16 to the membrane is adopting a fully extended state..101

Table 1. Summary of calculated 50% effective doses (ED50) for specific lysis in the in vitro ADCC assay and 50% inhibitory concentrations (IC50) for the in vitro neutralization assay...97

CHAPTER 1

Introduction

The discovery of Human Immunodeficiency Virus

On July 3, 1981, the *New York Times* published an article detailing the diagnosis of a rare form of cancer called Kaposi's Sarcoma among 41 homosexual men primarily in New York City and the San Francisco Bay Area over a period of 30 months (1). Until that time, the nationwide incidence of this disease was one case in every 1.5 million people as reported by the Centers for Disease Control (CDC). Two years later, with 1,350 reported cases, viral DNA from a human T-lymphotrophic retrovirus was isolated from the T cells of several patients exhibiting symptoms common to what is now called Acquired Immune Deficiency Syndrome (AIDS) (2). The enveloped virus was named Human Immunodeficiency Virus (HIV), and its discovery earned Drs. Françoise Barré-Sinoussi and Luc Montagnier the Nobel Prize in Physiology or Medicine in 2008. It was classified as belonging to genus *Lentivirus* in the *Retroviridae* family, which encode positive sense single-stranded RNA genomes. The lentiviral classification denotes the protracted period of latency which, in the case of HIV, can exceed seven years before the emergence of opportunistic infections that are collectively referred to as AIDS (3).

Today, HIV is a global pandemic. It is estimated that HIV/AIDS has claimed the lives of 25 million people with an estimated 33 million people living with the disease in 2007 (4). Although the development of life-extending highly active antiretroviral therapy (HAART) in 1996 is credited for the precipitous drop in AIDS-related morbidity and mortality in developed countries such as the United States (5), the costs associated with HAART and the lack of robust health care infrastructures has hindered their widespread use in most developing countries. For example, it was reported that by the end of 2007, the number of people receiving HAART in developing countries had finally reached the

target goal of 3 million – roughly 10% of those in need – two years after the 2005 deadline that WHO/UNAIDS had set in the "3 by 5" initiative proposed in 2003 (4). The combination of economic, logistical, and political barriers to treatment is particularly alarming given that nearly 95% of all HIV-infected people live in developing regions of the world such as sub-Saharan Africa and Asia (6). Education and sexually transmitted disease control programs have proven to be effective strategies for reducing the spread of HIV (7), but the need to develop an inexpensive vaccine capable of preventing or controlling an infection remains one of the most important challenges in combating the pandemic.

The modular nature of antibodies renders them amenable to the construction of antibody-like molecules with architectures that cannot be produced by the natural human immune system. Among others, these architectures include single-chain constructs encoding the variable heavy (V_H) and variable light (V_L) chains joined by short polypeptide linkers (scFv) (8), bispecific antibodies in which the two combining sites bear different specificities (9), and tetravalent bispecific molecules (10). In some cases, these formats have been shown to provide certain advantages over standard antibodies such as increased tissue penetration and more rapid pharmokinetics (11, 12), multispecificity, enhanced ability to recruit effector cells (13), and increased avidity (10).

The work described here was conducted over a five-year period in the laboratory of Pamela J. Björkman in collaboration with the laboratory of David Baltimore. The primary objective of this work was to utilize existing alternative antibodies architectures and design new ones to uncover principles underlying the mechanisms of HIV neutralization. For example, experiments were conducted to investigate whether cross-

linking epitopes is possible and/or important during neutralization and if other attributes of an antibody, such as size and flexibility, might influence neutralization potency. Attention was also given to the role of antibody-dependent cellular cytotoxicity (ADCC), an issue that has gained greater attention in recent years. Hopefully, the results described here will help guide future work towards the development of new therapeutics and/or vaccines, either in the form of a classical vaccine or an alternative approach such as gene therapy.

Origin and phylogeny of HIV

When compared to non-human primates (NHPs), the consensus sequence of HIV type 1 (HIV-1) appears to be most closely related to a strain of Simian Immunodeficiency Virus (SIV) that is endemic to *Pan troglodytes troglodytes (Ptt)*, a sub-species of chimpanzee (14). In the application of various methods of molecular clock analyses that utilize known mutation rates of HIV-1 to the comparison of sequences of RNA extracted from preserved human tissue samples dating between 1959 and 1960, current best estimates for the transfer of HIV-1 into the human population have converged on an initial date of approximately 1910 (15). The simplest and most widely accepted explanation for the crossover from *Ptt* to *Homo sapiens* is the "Hunter Theory", wherein a period of rapid population growth in sub-Saharan Africa in the early 20th century led to increased contact with and consumption of NHPs (16). The origin of the less virulent and

relatively rare HIV type 2 (HIV-2) is thought to share a similar history with respect to entry into the human population (17). More recently, the evaluation of blood samples collected in 2005 from hunters of NHPs in West Africa have identified new human T-lymphotropic retroviruses similar to HIV-1, suggesting that novel cross-species infection events are ongoing (18). Consequently, measures to reduce the hunting and/or consumption of NHPs may be critical to preventing the crossover of novel retroviruses into the human population.

Because HIV-1 currently accounts for nearly all HIV infections, the term HIV will be used to refer specifically to HIV-1 from this point forward. Among the three recognized phylogenetic branches of the virus, the main branch, group M, accounts for more then 99% of all infections worldwide (19). Distinct lineages within this branch are further subdivided into clades A though K on the basis of sequence diversity (20). Although diversity is found throughout the HIV genome, it is concentrated in the envelope (*Env*) gene, which encodes the gp160 precursor that is subsequently cleaved into gp120 and gp41, the two surface glycoproteins responsible for binding to target cells and catalyzing fusion.

The maximum level of sequence diversity observed between clades is approximately 35% when the HIV branch M-infected population is regarded as a whole, but sequence diversity may reach between 10% and 20% within one chronically infected individual (21, 22). This factor no doubt bears heavily on the ongoing failure to develop a prophylactic or therapeutic vaccine. Further diversity arises from epidemic circulating recombinant forms (CRFs) in which multiple clades are represented within a single isolate (23), presumably arising from recombination events following super-infections of

individuals who are already infected (24).

The global distribution of HIV clades is another factor that may prove critical to the design of a successful vaccine or therapy. Notwithstanding the impact that intercontinental travel has on the spread of HIV clades beyond their geographic origins, specific clades currently tend to be concentrated in well-defined regions. The most common clade today is C, which is primarily found in South Africa, China, and India; clade B is common to the Americas, Europe, Australia, and China; clade E, which is now formally classified a CRF of A and E, is found primarily in Southeast Asia; clades A and D are mostly found in East Africa, and nearly all clades circulate in West and Central Africa including the clades of the other two branches (24-26).

Though classification of HIV isolates by clade is based on general sequence diversity without specific attention to differences in epitopes thought to confer susceptibility to neutralization, at least one study suggests that the cross-reactivity of a particular antibody can be highly clade-specific. Using a small panel of monoclonal anti-HIV IgG subclass 1 antibodies, whose variable regions were derived from the B cells of clade B infected donors, researchers were able to show a statistically relevant clustering between branch M clades A through F on the basis of neutralization (19).

Natural history of HIV infection

In order to infect cells, HIV requires the presence of two receptors on the membranes of target cells: CD4, to which gp120 binds with low nanomolar affinity (27), and one of perhaps as many as a dozen G-protein-coupled receptors (GPCRs) (28, 29). Though CCR5, a principal HIV binding target, is expressed on multiple cell types

including T cells and the epithelial cells of the intestinal tract (30), early methods of characterizing HIV isolates based on the ability to infect different cell lines led researchers to divide them into two groups: R5 virus isolates, which preferentially bind CCR5 and infect activated $CD4^+$ memory T cells and macrophages (31, 32), and X4 isolates, which preferentially bind CXCR4 and infect naïve $CD4^+$ T cells (32, 33). In addition to CD4 and the dendritic cell (DC) C-type lectin receptors DC-SIGN and DC-SIGNR, these two co-receptors define the tropism of nearly all strains of HIV. Consequently, the primary cellular targets include dendritic cells, macrophages, and naïve, effector, and memory $CD4^+$ T cells (28). The microglial cells of the central nervous system expressing the CCR3 GPCR in addition to CCR5 constitute another target as their infection is thought to be an important step leading to the development of AIDS-related dementia by inducing the release of neurotoxins (34, 35).

When initially infected, the majority of people manifest flu-like symptoms and a high viral titer followed by activation of a cytotoxic T lymphocyte (CTL) response and, later, seroconversion (28). Following the early period of acute viremia, the concentration of virus drops to a lower so-called "set point," a leading prognostic indicator referring to the relatively stable viral titer observed after the early stage of acute viremia, and may remain relatively unchanged for many years without HAART (28). However, the peripheral $CD4^+$ T cell count will continue to decay throughout the period of clinical latency and, after dropping below approximately 500 cells/μL (28), viral concentration increases as AIDS-defining opportunistic infections emerge, ultimately resulting in death.

With the development of biological models for studying infection such as an *in vivo* model that pairs an HIV/SIV chimera (SHIV) challenge with rhesus macaques,

recent data are converging on the order and nature of early events that occur during mucosal transmission across the vaginal epithelium, the most common mode of HIV infection (reviewed in (36)). Following initial exposure, HIV is believed to penetrate several microns into the stratified squamous epithelium, probably by simple diffusion (37). In this tissue the virus encounters Langerhan cells, a type of DC that migrates within the epithelium and, because of the heavy glycosylation found on gp120, the virus binds to its receptors, DC-SIGN and DC-SIGNR (38-40). Following uptake of the virus, DCs have been observed *in vitro* to form filopodia in response to upregulation of Rho GTPase Cdc42, which then serve to efficiently transfer virus to CD4$^+$ T cells via an "infectious synapse" and initiate infection (41). Current evidence suggests that the preceding events probably occur within the first few hours of exposure (36).

After the first few days of infection, the rhesus macaque model revealed that a massive depletion of CD4$^+$ memory T cells surrounding the gut-associated lymphoid tissue (GALT) occurs concurrently with a rapid progression to peak viremia (42). The loss of CD4$^+$ T cells surrounding the GALT appears to persist in humans even after recovery of the circulating CD4$^+$ T cell population observed with HAART (36).

With the development of HAART, it has become possible to derive quantitative information about various steps in the natural history of infection. For example, upon administration of a protease inhibitor called Ritonavir (43) and a high-resolution time course of monitoring virus concentrations and CD4$^+$ T cell concentrations among HIV-infected patients, researchers were able to calculate that the average lifetime of a virus particle is ~7 hours, the average life-span of a productively-infected CD4$^+$ T cell is ~2 days, and up to ~10 billion new virus particles can be produced per day (44). With such

rapid turnover, the low fidelity of reverse transcriptase (45), and the prevalence of recombination (46), it is not surprising that resistance to individual inhibitors can sometimes be measured in days (47).

Interestingly, despite the sequence variation found in chronically infected individuals, the initial infection can almost always be traced back to a single founder virus whose sequence remains relatively unchanged throughout the initial infection period (48), and for unknown reasons, the co-receptor specificity of these founder viruses is usually for CCR5 rather than CXCR4. Selective pressure from the subsequent cytotoxic T lymphocyte (CTL) response then appears to induce the emergence of divergent virus populations (48).

Unfortunately, the development of potent autologously neutralizing antibodies (NAbs) remains conspicuously absent from the humoral immune responses of most individuals (49, 50). These observations and NHP studies in which depletion of CTLs (*i.e.,* $CD8^+$ T cells) led to uncontrolled SIV replication and accelerated death suggest that the early but temporary containment of viral replication that precedes the latency period is largely a function of the CTL response (49, 51, 52). The critical role that the CTL response plays in suppressing viral titer has fueled interest in the development of a CTL-based vaccine that targets conserved MHC class I epitopes but suffers from the fact that it does not address the putative requirement to neutralize cell-free virus particles, which is likely to prove critical to blocking the initial infection. Vaccine candidates have been tested in NHP challenge models (52-56), and the data support the contention that a pre-primed CTL response leads to better viral control (53), but a promising human CTL-based vaccine candidate has yet to emerge.

In contrast to the inherent limitation of a CTL vaccine, passively administered broadly cross-reactive monoclonal NAbs against HIV have been shown to enable NHPs to resist viral challenge as well as suppress viral rebound in humans following interruption of HAART (57-59). These results suggest that a properly designed immunogen capable of eliciting a potent NAb response or efforts to engineer more potent NAbs and antibody-like molecules could prevent an infection and possibly help control an established infection.

Structure of the HIV envelope spike and its mechanism of fusion

The *Env* gene encodes the glycoprotein gp160, which is cleaved into gp120 and gp41 by the cellular protease furin after passing through the Golgi complex (60). A high-resolution X-ray crystal structure of the entire trimeric gp120/gp41 assembly or the heterodimer has yet to be solved. However, sedimentation equilibrium analysis has shown that the mature fusion-competent viral spike present on the virion surface is a 442 kDa assembly of three copies of gp120 non-covalently associated with three copies of gp41 (61), and recent tomographic reconstructions of cryo-preserved HIV and SIV particles derived by electron microscopy have confirmed the three-fold symmetry of the spike (62-64). In addition, high resolution X-ray crystal structures of individual gp120 monomers in different conformational states and a portion of the six-helical bundle of gp41 representing its fusogenic conformation combined with extensive biochemical studies have contributed to the current model for fusion (65, 66).

Two of the high-resolution X-ray crystal structures of monomeric gp120 that are currently available provide substantial clarification to the question of why the human

immune system generally fails to develop potent NAbs against gp120 protein. The first crystal structure, solved in 1998 to 2.5 Å, is of a ternary complex of the hydrophobic core of gp120 derived from strain HxBc2 (also known as IIIB), a soluble form of CD4 composed of the first two domains (sCD4), and a Fab fragment of the modestly effective NAb 17b (67) (Fig. 1). In this ternary complex, nearly all glycosylation has been removed by mutation from gp120, which also carries single GlyAlaGly substitutions for 67 residues of the V1/V2 loops and 37 residues of the V3 loop, and is truncated at the N and C termini. The ability of this altered form of gp120 to still bind a panel of NAbs was confirmed (68).

The authors' analysis of the gp120 interface involved in binding CD4 revealed that the particularly recessed nature of the binding pocket may contribute to its ability to evade the humoral immune response as proposed in the "canyon hypothesis": a conserved binding site is hidden in a pocket that is sterically inaccessible to an antibody combining site while amino acid residues at the surface of the pocket are allowed to mutate with relative freedom (69). It was also observed that many of the contacts with CD4 are mediated by main chain instead of side chain atoms, leaving these residues relatively free to mutate under the selective pressure of the adaptive immune response.

Additional analysis of the interface with Nab 17b, which is known to overlap with the binding site for the chemokine coreceptor, revealed a "bridging sheet" composed of four anti-parallel β-strands, a result that was predicted from the conformational change associated with binding CD4 (27). This prediction appeared to have been validated when the structure of an uncomplexed, nearly fully glycosylated gp120 core from SIV was solved to 4 Å resolution (70) (Fig. 1). When comparing both gp120 crystal structures, it
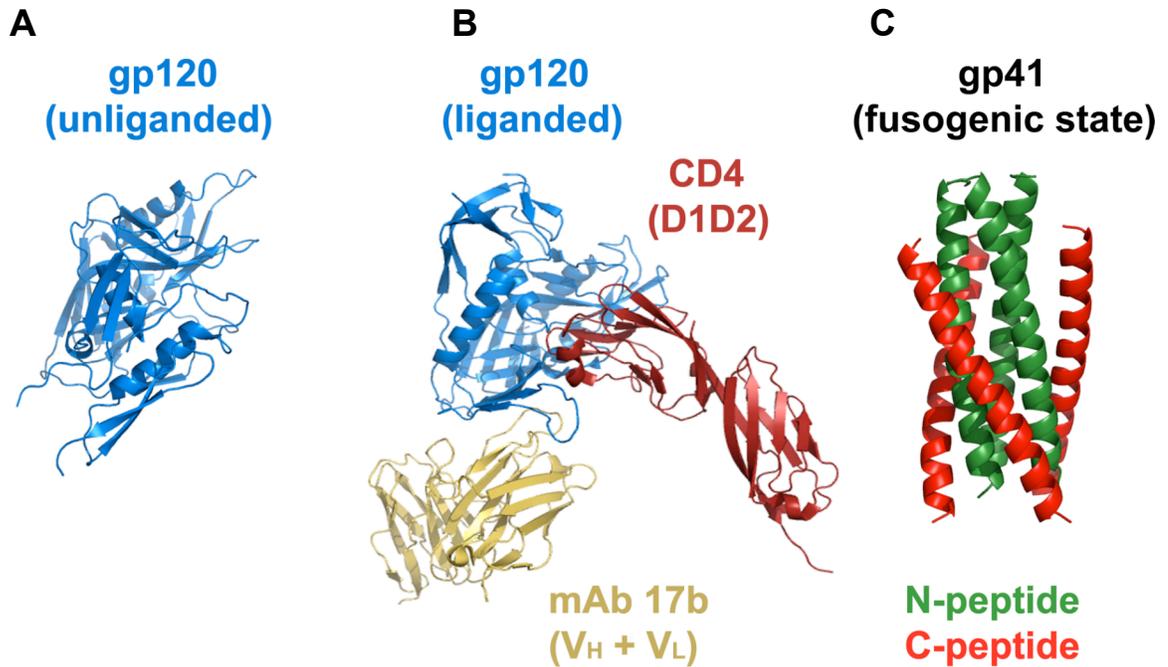
Figure 1. X-ray crystal structures of envelope spike proteins. (*A*) X-ray crystal structure of unliganded monomeric gp120 (60). (*B*) X-ray crystal structure of liganded monomeric gp120 bound to CD4 (domains 1 and 2), and mAb 17b (showing the V_H and V_L domains), which overlaps the co-receptor binding site (57). (*C*) X-ray crystal structure of N-peptides and C-peptides derived from gp41 in the six-helical bundle fusogenic state (71).

can be seen that in the unliganded structure the two β-strands belonging to the V1-V2 region are separated from the other two β-strands of the bridging sheet, β20 and β21, by nearly 20 Å, with most of the gap occupied by the amphipathic α1 helix. Packed against the β20/β21 portion of the bridging sheet is the now malformed CD4 binding loop. Consequently, the CD4 and co-receptor binding interfaces are nearly unrecognizable in this unliganded conformation. These observations are consistent with biochemical data supporting the notion that, concomitant with large enthalpic and entropic shifts derived from isothermal calorimetry measurements (27), gp120 undergoes an extensive conformational rearrangement in response to binding CD4, permitting assembly of the co-receptor binding site (27, 72-76).

Sequence analysis has shown that the gp120 subunit may be separated into five constant regions (C), which form the protein's core and participate in ligand-binding interactions, and five variable regions (V) responsible for shielding the core from immune detection, although studies have shown that V3 is a critical structural component of the chemokine coreceptor binding interface (77). Both gp120 and gp41 have conserved N-linked glycosylation sites; however, gp120 displays an especially high number of glycans, with carbohydrate contributing nearly half of its molecular weight – a feature that is believed to serve as a mechanism to reduce its antigenicity by shielding conserved protein epitopes and epitopes that are subject to selective pressure (78).

The mechanism of fusion is a multi-step process lasting between 15 and 30 minutes (79) for which many reagents exist that specifically target each of the various steps (Fig. 2). In the first step, gp120 engages the CD4 receptor, inducing a conformational change to expose the co-receptor binding site (27), a process that can be

blocked by reagents that target the CD4 binding site including small molecules, antibodies, and soluble forms of CD4 (80, 81). Following CD4 engagement, gp120 binds the co-receptor to form a stable attachment to the target cell, a step that can be blocked by small antibody fragments that can diffuse into the limited volume that separates the virus from the target cell (74).

After attachment, a fusion peptide at the N-terminus of gp41 is inserted into the target cell membrane to form what is called the pre-hairpin intermediate state (82), which can be targeted by peptides derived from the C-terminal heptad repeat region (HR2) of gp41 and by antibodies reactive to the N-terminal heptad repeat region (HR1) or the membrane proximal external region (MPER) (83-86). At some point during this process, the ability of gp120 to "shed" from the envelope spike is required for fusion to proceed (87). However, the precise point at which shedding takes place and whether this might present a rate-limiting step in the mechanism is not yet clear.

In the final step of the fusion mechanism, the three monomers of the C-terminal HR2 individually bind in a parallel orientation to the three grooves formed by the trimeric N-terminal HR1, forming a six-helical bundle or "hairpin" composed of both the N- and C-terminal portions of gp41 (65). This fusogenic conformation is comparable to observations made in structural studies of the envelope spikes of influenza type A and Moloney murine leukemia virus (88). It has further been demonstrated that peptides derived from the HR2 region exhibit a limited ability to partition into lipid environments (89, 90). Consequently, whether formation of the six-helical bundle creates the driving force for membrane fusion or simply represents the final conformation of the envelope spike following lipid mixing induced by these hydrophobic sequences is still debated (91, 92).
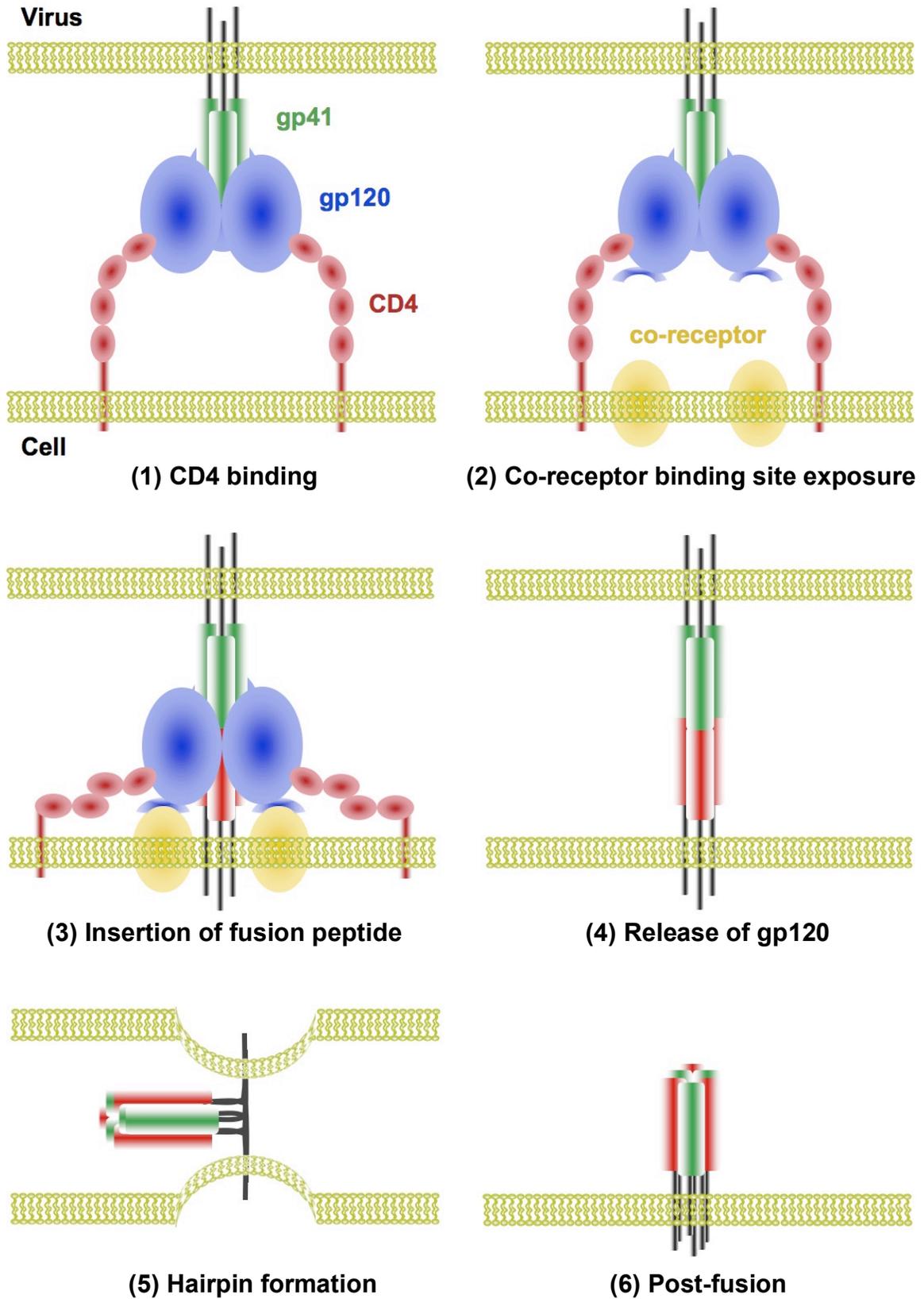
Figure 2. Schematic model of the HIV envelope spike fusion mechanism in 6 steps.

Isolation and characterization of broadly neutralizing monoclonal antibodies

The envelope spike comprises the antigenic target on the virus and infected cells against which neutralizing antibodies are directed (Fig. 3). A limited group of anti-HIV human monoclonal NAbs that have been extensively characterized including 2G12, 4E10, and 2F5 were isolated from relatively healthy HIV infected patients in the early to mid 1990s (75, 93, 94). Three other NAbs belonging to this extensively characterized group are b12 and X5, which were selected by phage display libraries derived from the bone marrow of HIV-infected patients (95, 96). The epitopes recognized by these NAbs have been mapped to either gp120 or gp41 and can be classified as belonging to one of five distinct groups: (1) the MPER of the gp41 ectodomain (*e.g.,* 4E10, 2F5, and Z13); (2) the CD4-binding site (CDbs) of gp120 (*e.g.,* b12); (3) the CD4-induced (CD4i) coreceptor binding site of gp120 (*e.g.,* 17b and X5); (4) the heavily glycosylated and immunologically "silent" face of gp120 (*e.g.,* 2G12); (5) the V3 loop of gp120 (*e.g.,* 447-52D). Although initial experiments indicated that these antibodies bound their epitopes with low nanomolar (nM) affinities when tested against antigens derived from various CCR5- and CXCR4-specific laboratory-adapted HIV strains (74, 75, 96-102), subsequent characterization in pseudovirus neutralization assays demonstrated enormous variation in potency against an extensive panel of primary isolates (19). This finding highlighted the need to carefully examine the *in vitro* cross-reactivity of any potential therapeutic candidates.

Given that each of these broadly NAbs was isolated from a patient infected with a virus strain belonging to clade B, it is not surprising that these antibodies were found to effectively neutralize many if not most representatives of that clade (19). However,

A

gp41

gp120

CD4

Fab 17b

B

Fab b12

C

Fab 2G12

D

Fab 4E10

E

Fab 2F5
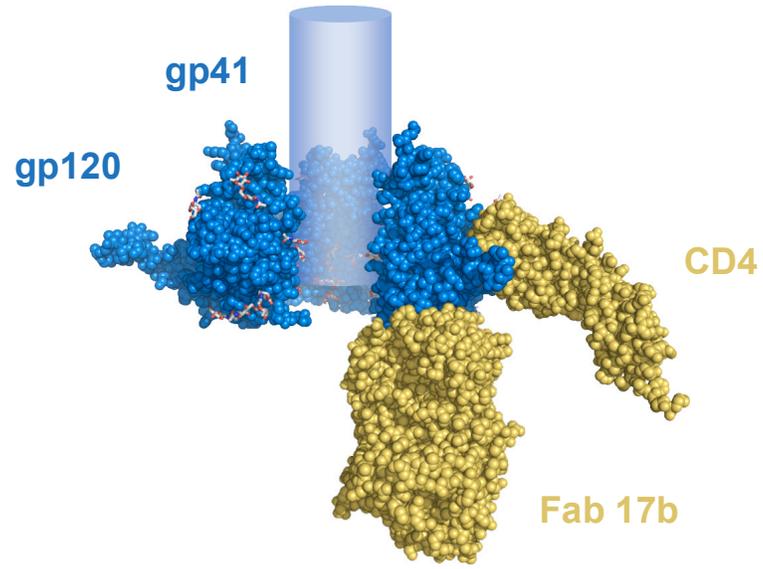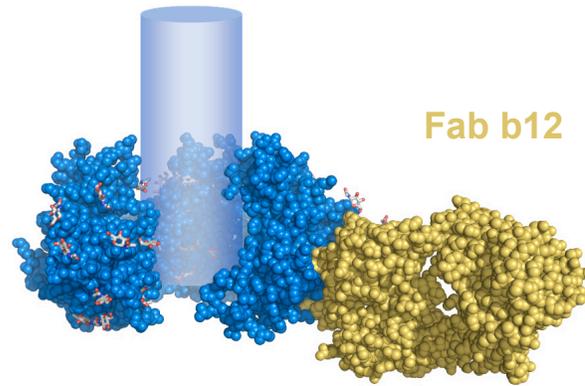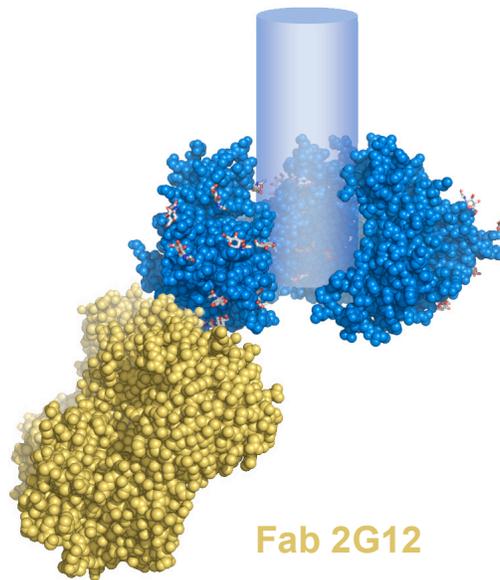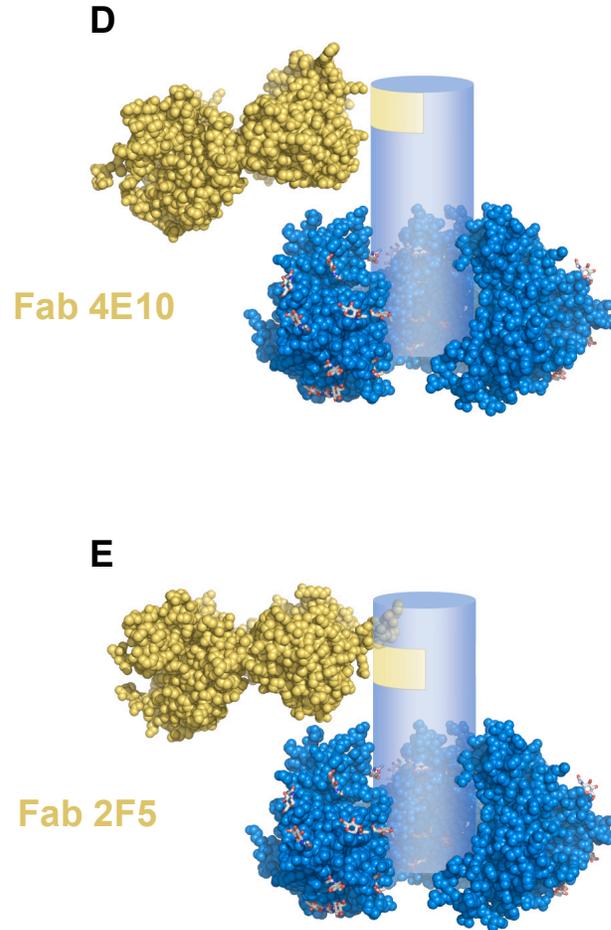
Figure 3. Models of the HIV envelope spike bound to CD4 and monoclonal antibodies. (*A*) CD4 (domains 1 and 2) and the CD4-induced mAb 17b (67), (*B*) the CD4-binding site mAb b12 (103), (*C*) the anti-carbohydrate antibody 2G12 (approximation of binding location) (104), (*D*) the anti-MPER antibody 4E10 (approximation of binding location) (105), and (*E*) the anti-MPER antibody 2F5 (approximation of binding location) (106). Trimeric gp120 and a schematic representation of gp41 are shown in blue.

unlike the anti-gp120 NAbs, which showed marked restrictions in clade-specific neutralization, the anti-gp41 NAbs, 4E10 and 2F5, were able to neutralize representatives from most clades with the exception of C and D for 2F5. The NAb X5 was exclusively restricted to clade B, 2G12 was primarily restricted to clades A and B, and b12 was primarily restricted to clades B, C, and D (19). The observation that anti-gp41 antibodies are by far the most cross-reactive may not be unexpected given that gp41 lacks the variable loops regions and the extensive glycosylation of gp120.

In spite of the limited success in identifying a small group of broadly reactive neutralizing antibodies (bNAbs), no one has yet been able to demonstrate a method to elicit them by vaccination, clear or effectively suppress a viral infection by their passive administration, or show that they can prevent infection using *in vivo* NHP models challenged with anything other than strains that are unusually sensitive to these antibodies. The following chapters include data that address the failure of these bNAbs to effectively suppress infection in humans and prevent infection in primates at realistically achievable concentrations. Findings related to steric occlusion of conserved binding sites and the impacts of antibody size and valency are presented in Chapter 2, limitations to epitope cross-linking are presented as a novel hypothesis called the "island effect" in Chapter 3, and deficiencies in the capacity of these bNAbs to trigger antibody-dependent cellular cytotoxicity (ADCC) are presented in Chapter 4.

References

1.   Altman LK (July 3, 1981) Rare Cancer Seen in 41 Homosexuals. The New York Times.

2.   Barre-Sinoussi F, *et al*. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220(4599):868-871.

3.   Berkelman RL, Heyward WL, Stehr-Green JK, & Curran JW (1989) Epidemiology of human immunodeficiency virus infection and acquired immunodeficiency syndrome. *Am J Med* 86(6):761-770.

4.   UNAIDS (2008) Report on the global AIDS epidemic. Geneva).

5.   Palella FJ, Jr.*, et al*. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* 338(13):853-860.

6.   UNAIDS/WHO (2007) AIDS Epidemic Update. in *Joint United Nations Programme on HIV/AIDS (UNAIDS) and World Health Organization (WHO))*.

7.   Grosskurth H, *et al*. (1995) Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. *Lancet* 346(8974):530-536.

8.   Bird RE*, et al*. (1988) Single-chain antigen-binding proteins. *Science* 242(4877):423-426.

9.   Staerz UD, Kanagawa O, & Bevan MJ (1985) Hybrid antibodies can target sites for attack by T cells. *Nature* 314(6012):628-631.

10. Zuo Z, Jimenez X, Witte L, & Zhu Z (2000) An efficient route to the production of an IgG-like bispecific antibody. *Protein Eng* 13(5):361-367.

11. Das D, Kriangkum J, Nagata LP, Fulton RE, & Suresh MR (2004) Development of a biotin mimic tagged ScFv antibody against western equine encephalitis virus: bacterial expression and refolding. *J Virol Methods* 117(2):169-177.

12. Wu AM, *et al.* (1996) Tumor localization of anti-CEA single-chain Fvs: improved targeting by non-covalent dimers. *Immunotechnology* 2(1):21-36.

13. Peipp M & Valerius T (2002) Bispecific antibodies targeting cancer cells. *Biochem Soc Trans* 30(4):507-511.

14. Keele BF, *et al.* (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 313(5786):523-526.

15. Worobey M, *et al.* (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455(7213):661-664.

16. Chitnis A, Rawls D, & Moore J (2000) Origin of HIV type 1 in colonial French Equatorial Africa? *AIDS Res Hum Retroviruses* 16(1):5-8.

17. Reeves JD & Doms RW (2002) Human immunodeficiency virus type 2. *J Gen Virol* 83(Pt 6):1253-1265.

18. Wolfe ND, *et al.* (2005) Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc Natl Acad Sci U S A* 102(22):7994-7999.

19. Binley JM, *et al.* (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78(23):13232-13252.

20. Kuiken C, *et al*. eds (1999) *A compilation and analysis of nucleic acid and amino acid sequences. In Human Retroviruses and AIDS*. (Los Alamos National Laboratory, Los Alamos, New Mexico).

21. Kuiken CL*, et al*. (1996) Evidence for limited within-person evolution of the V3 domain of the HIV-1 envelope in the amsterdam population. *Aids* 10(1):31-37.

22. Wain-Hobson S (1993) The fastest genome evolution ever described: HIV variation in situ. *Curr Opin Genet Dev* 3(6):878-883.

23. Robertson DL*, et al*. (2000) HIV-1 nomenclature proposal. *Science* 288(5463):55-56.

24. Burke DS (1997) Recombination in HIV: an important viral evolutionary strategy. *Emerg Infect Dis* 3(3):253-259.

25. Kuiken C, Thakallapalli R, Esklid A, & de Ronde A (2000) Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. *Am J Epidemiol* 152(9):814-822.

26. Taylor BS & Hammer SM (2008) The challenge of HIV-1 subtype diversity. *N Engl J Med* 359(18):1965-1966.

27. Myszka DG, *et al*. (2000) Energetics of the HIV gp120-CD4 binding reaction. *Proc Natl Acad Sci U S A* 97(16):9026-9031.

28. Janeway CA, Travers P, Walport M, & Schlomchik MJ (2005) *Immunobiology* (Garland Science Publishing, New York, NY) 6th Ed.

29. Shimizu N*, et al*. (1999) An orphan G protein-coupled receptor, GPR1, acts as a coreceptor to allow replication of human immunodeficiency virus types 1 and 2 in brain-derived cells. *J Virol* 73(6):5231-5239.

30. Meng G, *et al*. (2002) Primary intestinal epithelial cells selectively transfer R5 HIV-1 to CCR5+ cells. *Nat Med* 8(2):150-156.

31. Deng H, *et al*. (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 381(6584):661-666.

32. Ostrowski MA, *et al*. (1998) Expression of chemokine receptors CXCR4 and CCR5 in HIV-1-infected and uninfected individuals. *J Immunol* 161(6):3195-3201.

33. Feng Y, Broder CC, Kennedy PE, & Berger EA (1996) HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272(5263):872-877.

34. He J, *et al*. (1997) CCR3 and CCR5 are co-receptors for HIV-1 infection of microglia. *Nature* 385(6617):645-649.

35. Rock RB, *et al*. (2004) Role of microglia in central nervous system infections. (Translated from eng) *Clin Microbiol Rev* 17(4):942-964.

36. Boggiano C & Littman DR (2007) HIV's vagina travelogue. *Immunity* 26(2):145-147.

37. Hope TJ (2009) Interaction of HIV with Cells within Intact Mucosal Tissue. *HIV Immunobiology: From Infection to Immune Control (X4)*, (Keystone Symposia).

38. Mitchell DA, Fadden AJ, & Drickamer K (2001) A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *J Biol Chem* 276(31):28939-28945.

39. van Kooyk Y & Geijtenbeek TB (2003) DC-SIGN: escape mechanism for

pathogens. *Nat Rev Immunol* 3(9):697-709.

40. Guo Y*, et al.* (2004) Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* 11(7):591-598.

41. Nikolic DS, Lehmann M, Garcia E, Blanchet FP, & Piguet V (2009) Mechanisms of DC-T Cell HIV-1 Transmission via an Infectious Synapse. *Prevention of HIV/AIDS (X3)*, (Keystone Symposia).

42. Li Q*, et al.* (2005) Peak SIV replication in resting memory CD4+ T cells depletes gut lamina propria CD4+ T cells. *Nature* 434(7037):1148-1152.

43. Kempf DJ*, et al.* (1995) ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. (Translated from eng) *Proc Natl Acad Sci U S A* 92(7):2484-2488.

44. Perelson AS, Neumann AU, Markowitz M, Leonard JM, & Ho DD (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271(5255):1582-1586.

45. Roberts JD, Bebenek K, & Kunkel TA (1988) The accuracy of reverse transcriptase from HIV-1. *Science* 242(4882):1171-1173.

46. Charpentier C, Nora T, Tenaillon O, Clavel F, & Hance AJ (2006) Extensive recombination among human immunodeficiency virus type 1 quasispecies makes an important contribution to viral diversity in individual patients. *J Virol* 80(5):2472-2482.

47. Wei X*, et al.* (2002) Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents*

Chemother 46(6):1896-1905.

48.   Keele BF, *et al*. (2008) Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* 105(21):7552-7557.

49.   Koup RA, *et al*. (1994) Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J Virol* 68(7):4650-4655.

50.   Letvin NL (2004) Progress Toward an HIV Vaccine. *Annu Rev Med*.

51.   Schmitz JE, *et al*. (1999) Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science* 283(5403):857-860.

52.   McMichael AJ & Rowland-Jones SL (2001) Cellular immune responses to HIV. *Nature* 410(6831):980-987.

53.   Zhao J, *et al*. (2003) Improved protection of rhesus macaques against intrarectal simian immunodeficiency virus SIV(mac251) challenge by a replication-competent Ad5hr-SIVenv/rev and Ad5hr-SIVgag recombinant priming/gp120 boosting regimen. *J Virol* 77(15):8354-8365.

54.   Buge SL, *et al*. (1997) An adenovirus-simian immunodeficiency virus env vaccine elicits humoral, cellular, and mucosal immune responses in rhesus macaques and decreases viral burden following vaginal challenge. *J Virol* 71(11):8531-8541.

55.   Buge SL, *et al*. (1999) Factors associated with slow disease progression in macaques immunized with an adenovirus-simian immunodeficiency virus (SIV) envelope priming-gp120 boosting regimen and challenged vaginally with

SIVmac251. *J Virol* 73(9):7430-7440.

56.   Zwick MB, *et al*. (2001) Neutralization synergy of human immunodeficiency virus type 1 primary isolates by cocktails of broadly neutralizing antibodies. *J Virol* 75(24):12198-12208.

57.   Emini EA, *et al*. (1992) Prevention of HIV-1 infection in chimpanzees by gp120 V3 domain-specific monoclonal antibody. *Nature* 355(6362):728-730.

58.   Parren PW, *et al*. (2001) Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. *J Virol* 75(17):8340-8347.

59.   Trkola A, *et al*. (2005) Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat Med*. 11(6):615-622.

60.   Peterlin MB & Trono D (2003) Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat Rev Immunol* 3(2):97-107.

61.   Center RJ, *et al*. (2002) Oligomeric structure of the human immunodeficiency virus type 1 envelope protein on the virion surface. *J Virol* 76(15):7863-7867.

62.   Zhu P, *et al*. (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441(7095):847-852.

63.   Zanetti G, Briggs JA, Grunewald K, Sattentau QJ, & Fuller SD (2006) Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS Pathog* 2(8):e83.

64.   Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, & Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455(7209):109-113.

65.   Chan DC & Kim PS (1998) HIV entry and its inhibition. *Cell* 93(5):681-684.

66.   Doms RW & Trono D (2000) The plasma membrane as a combat zone in the HIV battlefield. *Genes Dev* 14(21):2677-2688.

67.   Kwong PD, *et al*. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393(6686):648-659.

68.   Binley JM, *et al*. (1998) Analysis of the interaction of antibodies with a conserved enzymatically deglycosylated core of the HIV type 1 envelope glycoprotein 120. *AIDS Res Hum Retroviruses* 14(3):191-198.

69.   Rossmann MG, *et al*. (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* 317(6033):145-153.

70.   Chen B, *et al*. (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* 433(7028):834-841.

71.   Chan DC, Fass D, Berger JM, & Kim PS (1997) Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89(2):263-273.

72.   Sullivan N, *et al*. (1998) CD4-Induced conformational changes in the human immunodeficiency virus type 1 gp120 glycoprotein: consequences for virus entry and neutralization. *J Virol* 72(6):4694-4703.

73.   Dey B, Del Castillo CS, & Berger EA (2003) Neutralization of human immunodeficiency virus type 1 by sCD4-17b, a single-chain chimeric protein, based on sequential interaction of gp120 with CD4 and coreceptor. *J Virol* 77(5):2859-2865.

74.   Labrijn AF, *et al*. (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(19):10557-10565.

75.   Thali M, *et al*. (1993) Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J Virol* 67(7):3978-3988.

76.   Wu L, *et al*. (1996) CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* 384(6605):179-183.

77.   Starcich BR, *et al*. (1986) Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 45(5):637-648.

78.   Wei X, *et al*. (2003) Antibody neutralization and escape by HIV-1. *Nature* 422(6929):307-312.

79.   Dimitrov AS, *et al*. (2007) Exposure of the membrane-proximal external region of HIV-1 gp41 in the course of HIV-1 envelope glycoprotein-mediated fusion. *Biochemistry* 46(5):1398-1401.

80.   Briz V, Poveda E, & Soriano V (2006) HIV entry inhibitors: mechanisms of action and resistance pathways. *J Antimicrob Chemother* 57(4):619-627.

81.   Burton DR, Stanfield RL, & Wilson IA (2005) Antibody vs. HIV in a clash of evolutionary titans. *Proc Natl Acad Sci U S A* 102(42):14943-14948.

82.   Harrison SC (2005) Mechanism of membrane fusion by viral envelope proteins. *Adv Virus Res* 64:231-261.

83.   Jiang S, Lin K, Strick N, & Neurath AR (1993) HIV-1 inhibition by a peptide. *Nature* 365(6442):113.

84. Wild CT, Shugars DC, Greenwell TK, McDanal CB, & Matthews TJ (1994) Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proc Natl Acad Sci U S A* 91(21):9770-9774.

85. Luftig MA*, et al.* (2006) Structural basis for HIV-1 neutralization by a gp41 fusion intermediate-directed antibody. *Nat Struct Mol Biol* 13(8):740-747.

86. Frey G*, et al.* (2008) A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. *Proc Natl Acad Sci U S A* 105(10):3739-3744.

87. Binley JM*, et al.* (2003) Redox-triggered infection by disulfide-shackled human immunodeficiency virus type 1 pseudovirions. *J Virol* 77(10):5678-5684.

88. Weissenhorn W, Dessen A, Harrison SC, Skehel JJ, & Wiley DC (1997) Atomic structure of the ectodomain from HIV-1 gp41. *Nature* 387(6631):426-430.

89. Sanchez-Martinez S, Lorizate M, Katinger H, Kunert R, & Nieva JL (2006) Membrane association and epitope recognition by HIV-1 neutralizing anti-gp41 2F5 and 4E10 antibodies. *AIDS Res Hum Retroviruses* 22(10):998-1006.

90. Sun ZY*, et al.* (2008) HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane. *Immunity* 28(1):52-63.

91. Barbato G*, et al.* (2003) Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. *J Mol Biol* 330(5):1101-1115.

92. Sackett K, Nethercott M, Shai Y, & Weliky D (2009) Hairpin folding of HIV gp41 abrogates lipid mixing function at physiologic pH and inhibits lipid mixing

by exposed gp41 constructs. *Biochemistry*. Feb 17 [Epub ahead of print]

93.   Robinson JE, Holton D, Pacheco-Morell S, Liu J, & McMurdo H (1990) Identification of conserved and variant epitopes of human immunodeficiency virus type 1 (HIV-1) gp120 by human monoclonal antibodies produced by EBV-transformed cell lines. *AIDS Res Hum Retroviruses* 6(5):567-579.

94.   Burton DR, *et al*. (2004) HIV vaccine design and the neutralizing antibody problem. *Nat Immunol* 5(3):233-236.

95.   Moulard M, *et al*. (2002) Broadly cross-reactive HIV-1-neutralizing human monoclonal Fab selected for binding to gp120-CD4-CCR5 complexes. *Proc Natl Acad Sci U S A* 99(10):6913-6918.

96.   Zwick MB, *et al*. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J Virol* 75(22):10892-10905.

97.   Pantophlet R, Wilson IA, & Burton DR (2004) Improved design of an antigen with enhanced specificity for the broadly HIV-neutralizing antibody b12. *Protein Eng Des Sel* 17(10):749-758.

98.   Hoffman TL, *et al*. (1999) Stable exposure of the coreceptor-binding site in a CD4-independent HIV-1 envelope protein. *Proc Natl Acad Sci U S A* 96(11):6359-6364.

99.   Wyatt R, *et al*. (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(9):5723-5733.

100.   Darbha R, *et al*. (2004) Crystal structure of the broadly cross-reactive HIV-1-

neutralizing Fab X5 and fine mapping of its epitope. *Biochemistry* 43(6):1410-1417.

101.   Scanlan CN, *et al.* (2002) The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. *J Virol* 76(14):7306-7321.

102.   Mo H, *et al.* (1997) Human immunodeficiency virus type 1 mutants that escape neutralization by human monoclonal antibody IgG1b12. off. *J Virol* 71(9):6869-6874.

103.   Zhou T, *et al.* (2007) Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* 445(7129):732-737.

104.   Calarese DA, *et al.* (2003) Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 300(5628):2065-2071.

105.   Cardoso RM, *et al.* (2005) Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity* 22(2):163-173.

106.   Ofek G, *et al.* (2004) Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol* 78(19):10724-10737.

CHAPTER 2

Examination of the contributions of size and avidity to the

neutralization mechanisms of the anti-HIV antibodies b12 and 4E10

Examination of the contributions of size and avidity to the neutralization mechanisms of the anti-HIV antibodies b12 and 4E10

Joshua S. Klein[a], Priyanthi N. P. Gnanapragasam[b], Rachel P. Galimidi[b], Christopher P. Foglesong[b], Anthony P. West, Jr.[b], and Pamela J. Bjorkman[b,c,1]
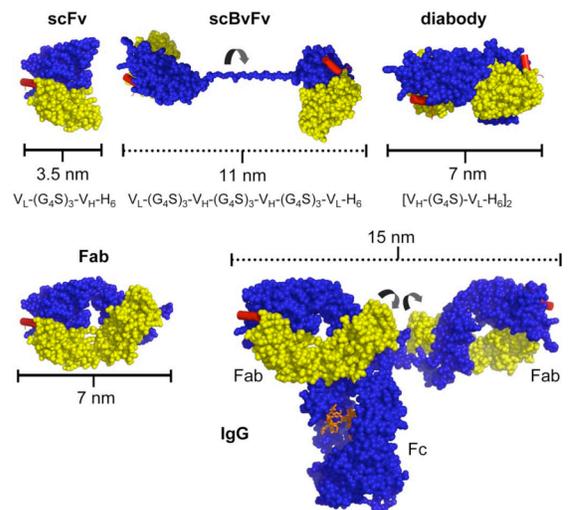
[a]Graduate option in Biochemistry and Molecular Biophysics, [b]Division of Biology, [c]Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125

Monoclonal antibodies b12 and 4E10 are broadly neutralizing against a variety of strains of the human immunodeficiency virus type 1 (HIV-1). The epitope for b12 maps to the CD4-binding site in the gp120 subunit of HIV-1's trimeric gp120-gp41 envelope spike, whereas 4E10 recognizes the membrane-proximal external region (MPER) of gp41. Here, we constructed and compared a series of architectures for the b12 and 4E10 combining sites that differed in size, valency, and flexibility. In a comparative analysis of the ability of the b12 and 4E10 constructs to neutralize a panel of clade B HIV-1 strains, we observed that the ability of bivalent constructs to cross-link envelope spikes on the virion surface made a greater contribution to neutralization by b12 than by 4E10. Increased distance and flexibility between antibody combining sites correlated with enhanced neutralization for both antibodies, suggesting restricted mobility for the trimeric spikes embedded in the virion surface. The size of a construct did not appear to be correlated with neutralization potency for b12, but larger 4E10 constructs exhibited a steric occlusion effect, which we interpret as evidence for restricted access to its gp41 epitope. The combination of limited avidity and steric occlusion suggests a mechanism for evading neutralization by antibodies that target epitopes in the highly conserved MPER of gp41.

Fig. 1. Structures of antibody constructs. Space-filling models are presented above a description of the domain organization for each construct (V_L, variable light; V_H, variable heavy; (G_4S), Gly-Ser linker; H_6, 6×-His tag). Models were constructed by using coordinates for the heavy (blue) and light (yellow) chains of Fab 4E10 and its peptide epitope (red) (PDB ID code 1TZG) (34). For the diabody model, 2 4E10 V_H–V_L pairs were aligned to the structure of diabody L5MK16 (PDB ID code 1LMK) (30). For the IgG model, 2 4E10 Fabs were used to replace the b12 Fabs in the structure of intact IgG1 b12 (PDB ID code 1HZH) (55). Solid lines indicate approximate dimensions for the scFv, diabody, and Fab. Dotted lines indicate approximate maximal distances between combining sites for the scBvFv and IgG. Curved black arrows indicate axes of rotation.

HIV type 1 (HIV-1) is an enveloped virus that presents severe challenges to eliciting effective antibody-mediated immune responses because it employs multiple strategies to evade antibodies. The virus rapidly mutates to change residues on its surface (1), conceals other potential antibody epitopes with carbohydrates (2), hides conserved regions at interfaces by oligomerization, and prevents access to conserved regions by conformational masking and steric occlusion (2–5). Despite these escape mechanisms, a limited number of broadly neutralizing antibodies have been isolated from HIV-1-infected individuals over the past few decades (reviewed in ref. 6). They target well-defined epitopes on both subunits of the HIV-1 envelope spike, a trimeric complex composed of 3 copies of 2 noncovalently associated glycoproteins, gp120 and gp41. One such antibody called b12 binds to an epitope that overlaps the host receptor (CD4)-binding site on gp120 (7, 8), and another called 4E10 binds to an epitope in the highly conserved membrane proximal external region (MPER) of gp41 (9–12). Both antibodies were shown to be broadly neutralizing across a diverse panel of HIV-1 strains, although 4E10 exceeded b12 in the breadth of its reactivity (13).

The neutralization potency of an antibody against a virus can be improved by orders of magnitude through the effects of avidity (14–18). The term avidity in the context of antibodies refers to their ability to simultaneously bind 2 physically linked antigens (e.g., 2 spikes on the surface of the same virus) by using the 2 identical combining sites located at the tips of their Fab (antigen-binding fragment) arms (19) (Fig. 1). In order for avidity to occur, the antigen sites must be present at sufficient density such that once the first Fab has bound, the second Fab

can bind its partner before the first Fab dissociates. The number of spikes on HIV-1 is ≈15 per virion (20–23), whereas ≈450 spikes per virion have been observed on the similarly sized influenza type A virus (24). The extent to which the relatively low density of HIV-1 envelope spikes might impact the avidity of anti-HIV-1 antibodies is not yet understood.

Our objective in the present study was to ask how the difference between monovalence and bivalence coupled with
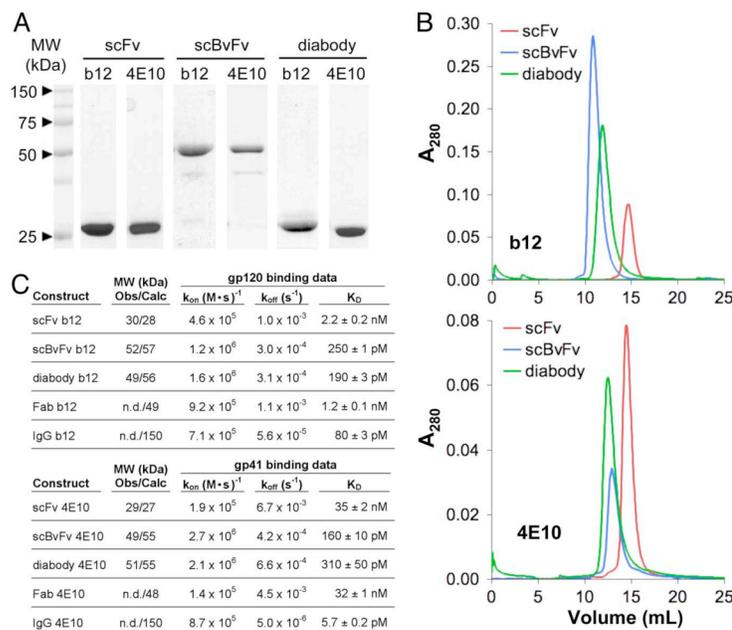
Fig. 2. Biophysical characterization of the antibody constructs. (*A*) Reduced SDS/PAGE. (*B*) Gel filtration profiles. (*C*) Molecular weight determinations and binding experiments. Observed results from static light scattering experiments (Obs) are presented beside molecular weights calculated from the relevant sequence (Calc) in column 2 (n.d., not done). Kinetic and equilibrium constants are presented in columns 3–5.

differences in size and flexibility contribute to the neutralization mechanisms of b12 and 4E10. Using an in vitro neutralization assay, we compared the potencies of b12 and 4E10 constructs against a panel of clade B HIV-1 strains. Our results demonstrated that avidity enhanced neutralization by IgG b12 but only weakly enhanced neutralization by IgG 4E10, and the contribution of avidity to b12-mediated neutralization was usually most apparent for strains that were relatively insensitive to monovalent b12 reagents. Moreover, we observed that flexibility and distance between the antigen-binding sites of bivalent forms of both antibodies enhanced neutralization potency and that increased size limited neutralization by 4E10 but not b12. The implications of these results on antibody escape by HIV-1 and vaccine design are discussed.

Results

Neutralizing Antibody Fragments Are Stable and Exhibit Correct Oligomerization. To systematically compare affinities and neutralization potencies as a function of size, number, and arrangement of combining sites, we produced monovalent and bivalent forms of b12 and 4E10. As monovalent forms, we produced the combining sites as Fabs and as scFvs (single chain variable fragments), in which a 15-residue flexible Gly-Ser linker was used to link the variable heavy and variable light (V_H and V_L) domains in a single polypeptide (25, 26) (Fig. 1). We made 3 different bivalent forms of each antibody: the traditional IgG, a single chain bivalent Fv (scBvFv), and a diabody (Fig. 1). The scBvFv was constructed by joining 2 scFv fragments with a third Gly-Ser linker, thereby forming a single polypeptide chain with 2 antibody combining sites of identical specificities (27). In this form of bivalent reagent, the 2 combining sites are expected to be free to rotate with respect to each other. The diabody form was constructed by expressing a scFv with a short linking region (28), which promoted pairing between a V_H domain and a V_L

domain on separate polypeptides to form a 3D domain-swapped dimer (29). Relative to scBvFvs, diabodies are expected to be more rigid, with 2 combining sites facing in approximately opposite directions (30). In total, we produced the Fab and IgG forms of b12 and 4E10 as well as 6 different scFv-based constructs (scFv b12, scBvFv b12, diabody b12, scFv 4E10, scBvFv 4E10, and diabody 4E10).

The scFv-based proteins were purified by Ni-NTA and size-exclusion chromatography and analyzed by SDS/PAGE (Fig. 2*A* and *B*). To verify that each of the proteins exhibited the expected oligomeric state, molecular weights were determined by in-line multiangle static light scattering coupled with size-exclusion chromatography (Fig. 2*C*). The results were consistent with the theoretical molecular weights of the scBvFvs and diabodies, which are approximately twice the molecular weight of a scFv, demonstrating that the scFvs and scBvFvs were monomeric and the diabodies were dimeric.

Bivalent b12 and 4E10 Reagents Can Bind with Avidity. The antigen binding activities of the b12 and 4E10 proteins were evaluated by using a surface plasmon resonance-based binding assay. For these experiments, we injected b12 reagents over immobilized monomeric gp120 and 4E10 reagents over immobilized gp41. During neutralization, antibodies bind to a gp120-gp41 envelope spike trimer on the surface of the virus rather than to the separated chains that can be expressed and purified for binding assays. Thus, an affinity derived from this binding assay cannot be used to deduce the affinity of an antibody for its epitope on the surface of a virus. Instead, the binding assays were used to verify that each of the reagents bound its antigen and to determine whether the bivalent constructs could cross-link immobilized antigens on the sensor surface, which would be revealed by avidity effects resulting in higher apparent affinities relative to the counterpart monovalent constructs.

Table 1. Strain-specific IC$_{50}$ neutralization values (nM) for each antibody construct

	Antibody construct									
	b12					4E10				
Virus strain	scFv	scBvFv	diabody	Fab	IgG	scFv	scBvFv	diabody	Fab	IgG
6535.3	760 ± 280	41 ± 6	1,100 ± 200	260 ± 80	19 ± 3	14 ± 3	19 ± 3	220 ± 90	34 ± 8	5.0 ± 1.3
QH0692.42	110 ± 10	31 ± 3	48 ± 10	76 ± 10	5.6 ± 0.4	98 ± 12	130 ± 30	1,100 ± 200	220 ± 50	38 ± 4
RHPA4259.7	12 ± 1	7.5 ± 1.2	19 ± 2	11 ± 2	1.0 ± 0.2	100 ± 20	320 ± 40	680 ± 110	760 ± 200	250 ± 50
SC422661.8	57 ± 8	12 ± 2	15 ± 4	28 ± 7	3.8 ± 0.5	10 ± 1	15 ± 2	54 ± 9	28 ± 3	15 ± 3
SF162	4.9 ± 1.7	1.2 ± 0.4	4.6 ± 0.2	2.6 ± 1.0	0.26 ± 0.04	56 ± 23	90 ± 25	480 ± 190	290 ± 130	19 ± 8
THRO4156.18	110 ± 20	12 ± 2	62 ± 29	130 ± 30	4.3 ± 1.4	35 ± 3	9.3 ± 1.5	84 ± 23	42 ± 10	3.4 ± 0.8
REJO4541.67	810 ± 250	33 ± 7	680 ± 140	290 ± 100	10 ± 2	—	—	—	—	—
WITO4160.33	390 ± 120	250 ± 70	n.d.	320 ± 60	22 ± 3	—	—	—	—	—
TRJO4551.58	—	—	—	—	—	92 ± 27	170 ± 30	610 ± 130	240 ± 30	35 ± 7
TRO.11	—	—	—	—	—	8.4 ± 0.9	27 ± 3	200 ± 50	19 ± 5	3.2 ± 0.6

n.d., not done. See discussion in *SI Text*.

All of the b12 and 4E10 reagents exhibited high antigen binding affinities with equilibrium dissociation constants (K_{DS}) in the nanomolar or picomolar range [Fig. 2C and supporting information (SI) Fig. S1]. The K_{DS} for the monovalent b12 reagents (scFv and Fab) were in close agreement (2.2 nM and 1.2 nM, respectively). All of the bivalent b12 reagents bound to gp120 with higher apparent affinities: 80 pM for IgG b12, 250 pM for scBvFv b12, and 190 pM for diabody b12 (Fig. 2C), demonstrating that each of the bivalent constructs contained 2 functional antigen-binding sites that could cross-link adjacent immobilized antigens. The larger distance between the binding sites in an IgG compared with the binding sites in a scBvFv or diabody (Fig. 1) would be expected to lead to increased cross-linking efficiency, rationalizing the higher apparent affinity of the IgG compared with those of the scBvFv and diabody. The results obtained for the 4E10 reagents also showed an affinity enhancement for the bivalent reagents over the monovalent reagents: The scFv and Fab bound to gp41 with K_{DS} of 35 nM and 32 nM, respectively, consistent with the 20 nM K_D reported for Fab 4E10 binding to a gp41-derived peptide (31), and the IgG, scBvFv, and diabody bound with apparent K_{DS} of 5.7 pM, 160 pM, and 310 pM, respectively (Fig. 2C).

Ability to Cross-Link Epitopes on a Virus Contributes to Neutralization by b12. Pseudovirus neutralization assays were performed for antibody constructs against a panel of 10 primary virus strains from clade B (32). Eight were originally selected for evaluating IgG b12 and IgG 4E10, but we replaced TRJO4551.58 and TRO.11 with REJO4541.67 and WITO4160.33 in the b12 analyses because they were insensitive to all of the b12 reagents except IgG b12. From plots of inhibitor concentration versus percentage inhibition, we derived molar concentrations at which 50% inhibition was observed (IC$_{50}$ values) for each potential inhibitor (Table 1 and Fig. S2). We then compared various pairs of antibody architectures by calculating the ratio of their average molar IC$_{50}$ values across all strains (Fig. 3).

All of the b12 reagents neutralized b12-sensitive virus isolates, but the bivalent IgG and scBvFv constructs were more potent than the monovalent scFv and Fab forms: IgG b12 was an average of 34-fold more potent than scFv b12 (i.e., the average molar IC$_{50}$ value for the scFv divided by the average value for the IgG was 34) and 17-fold more potent than Fab b12 (Fig. 3), and scBvFv b12 was an average of 6.0-fold more potent than scFv b12 and 2.9-fold more potent than Fab b12 (Fig. 3). Diabody b12,
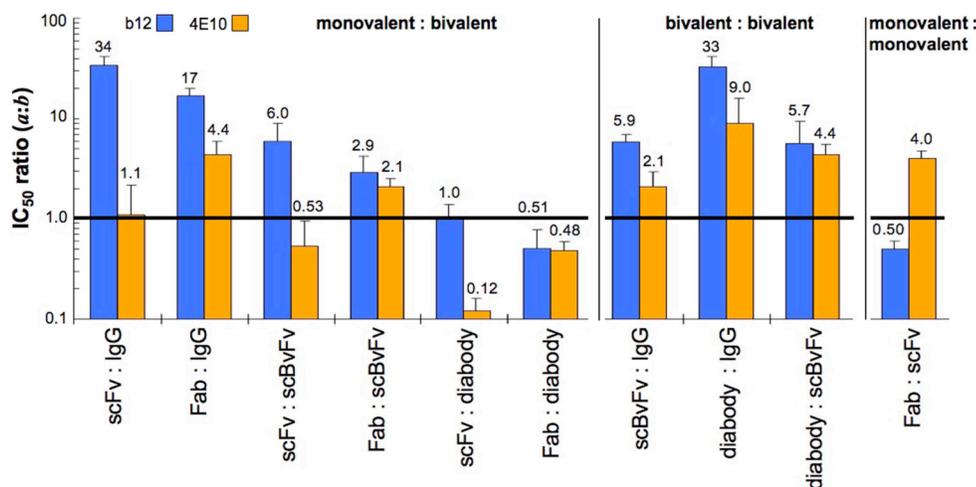
Fig. 3. Bar graph of ratios of average molar IC$_{50}$ values (arithmetic means) for b12 constructs (blue) and 4E10 constructs (orange). Reagent pairs with an average ratio of 1.0 (black line) are equal in average potencies. Ratios >1.0 indicate that reagent *b* is more potent than reagent *a*. Ratios <1.0 indicate that reagent *a* is more potent than reagent *b*. Error bars represent the standard errors calculated from the variability in strain-specific ratios for each pair of reagents.

however, was generally indistinguishable in neutralization potency when compared with the monovalent construct (scFv/diabody average IC_{50} ratio was 1.0 and the Fab/diabody ratio was 0.51).

The increased potencies of the IgG and scBvFv forms of b12 could result from their ability to cross-link epitopes on the surface of the virus (i.e., avidity), their larger sizes or different domain structures compared with the monovalent forms, or a combination of both. None of the monovalent constructs were as large as an IgG, so the effects of size and valency could not be separated in comparisons involving the bivalent IgG architecture. However, scBvFv b12, which contains 4 domains that are comparable in size and structure to the 4 domains of monovalent Fab b12, exhibited greater neutralization potency than the Fab for all strains tested, with an average increase of 2.9-fold (Table 1 and Fig. 3). The comparison between scBvFv b12 and scFv b12 allows us to control for potential effects of domains outside the variable regions impacting affinity and specificity in binding because both constructs contain only V_L and V_H domains, yet the bivalent scBvFv b12 showed an average 6.0-fold increase in neutralization potency compared with scFv b12. We suggest that the increased potencies of the bivalent IgG and scBvFv forms of b12 relates to their abilities to cross-link epitopes on the virus, with the larger distance between combining sites in the IgG compared with the scBvFv permitting more cross-linking. Diabody b12, although bivalent and able to cross-link immobilized gp120 in a binding assay (Figs. 1 and 2C), was equivalent to scFv b12 in neutralization potency (Fig. 3), suggesting that the relatively rigid pairing of 2 combining sites and shorter distance between combining sites did not permit efficient cross-linking on a viral surface. Comparing scBvFv b12 with the similarly sized diabody supports the conclusion that flexibility between the antibody combining sites is important for cross-linking epitopes on a virus in b12-mediated neutralization as the scBvFv exhibited a 5.7-fold average increased potency compared with the diabody.

Neutralization by 4E10 Involves only Minimal Cross-Linking of Epitopes on a Virus. A comparison of bivalent and monovalent 4E10 constructs shows that the 4E10 bivalent reagents exhibited only modest improvements in neutralization potency compared with monovalent constructs. For example, IgG 4E10 showed a 1.1-fold and 4.4-fold improvement in potency compared with the scFv and Fab, respectively, and the scBvFv was nearly equivalent to scFv 4E10 and only slightly more potent than Fab 4E10 (Fig. 3). These results suggest that IgG 4E10 has a minimal ability to cross-link epitopes on a virus and that the flexible scBvFv 4E10 generally behaved as a monovalent reagent, as evidenced by the 0.53 scFv/scBvFv average IC_{50} ratio for 4E10 versus a ratio of 6.0 for the comparable b12 reagents. As was observed for b12, diabody 4E10 showed no increase in potency compared with the monovalent reagents (scFv/diabody average IC_{50} ratio of 0.12 and Fab/diabody average IC_{50} ratio of 0.48), demonstrating that neither diabody could efficiently cross-link epitopes on the surface of a virus.

Neutralization Potencies Suggest a Size-Restricted Epitope for 4E10, but Not b12. Comparison of IC_{50} ratios indicates that smaller and/or more flexible 4E10 reagents were generally more potent in neutralization than larger and/or less flexible reagents, a relationship that was not observed for the b12 reagents. For example, scFv 4E10 was an average of 8.3-fold more potent than the larger diabody 4E10 (average scFv/diabody IC_{50} ratio of 0.12), whereas the scFv b12 and diabody b12 were equally potent (Fig. 3). In addition, scFv 4E10 was systematically more potent than the larger Fab 4E10 (an average 4.0-fold potency increase), contrasting with scFv b12, which was 2.0-fold less potent than Fab b12 (Fig. 3 and Table 1). Indeed, the IgG 4E10 was the only reagent larger than scFv 4E10 that was also more potent, but only in 6 of the 8 strains tested (Table 1 and Table S1); for strains RHPA4259.7 and SC422661.8 the scFv was more potent than the

IgG. By contrast, we observed no instances in which scFv b12 was as potent as IgG b12; the smallest difference in potencies was 12-fold with an average 34-fold difference (Fig. 3 and Table S1). These results are compatible with partial steric occlusion of the 4E10 epitope such that the larger 4E10 reagents are unable to gain complete access. Given that Fab 4E10 was found to be an average of 4.0-fold less potent than scFv 4E10, the occlusion appeared to be somewhat overcome in the case of the IgG by a modest ability to cross-link, thereby offsetting the steric penalty. In support of the hypothesis that flexibility in a 4E10 reagent improves access to the 4E10 epitope, we point to a comparison between diabody 4E10 and scBvFv 4E10, both of which were functioning as monovalent reagents during neutralization in the majority of strains tested (Table S1): Although similar in size, the more rigid diabody exhibited an average 4.4-fold weaker neutralization potency than the scBvFv. Taken together, the comparison of neutralization potencies for the IgG, Fab, diabody, and scFv forms of 4E10 and b12 suggested that the larger sizes of the IgG, Fab, scBvFv, and diabody forms of 4E10 prevented complete access to its epitope on gp41. See also Tables S2 and S3 for additional information.

Discussion

In this investigation, we asked whether alternative antibody architectures that do not naturally occur, such as a scFv, scBvFv, or diabody, could be used to further our understanding of the mechanisms by which the anti-gp120 antibody b12 and the anti-gp41 antibody 4E10 neutralize primary isolates in clade B of HIV-1. A comparative analysis of the neutralization potencies of these architectures as well as Fab and IgG forms of these antibodies yielded several conclusions that were consistent across multiple strains.

First, our analysis suggested that cross-linking HIV-1 epitopes contributes to the neutralization mechanism of IgG b12 but is less apparent for neutralization by IgG 4E10. Inefficient cross-linking by 4E10 may be related to its orientation when binding gp41: Previous reports suggested that a bound 4E10 Fab is oriented approximately perpendicular to the viral envelope (33, 34), which would require an I-shaped conformation of an IgG if both Fabs were simultaneously engaged. By contrast, b12 Fabs bind approximately parallel to the viral envelope (20), which could be achieved by a T- or Y-shaped conformation. However, although bivalency was more important for b12- than 4E10-dependent neutralization, the avidity-dependent increase in potency for b12 was limited relative to IgG/Fab comparisons of antibodies that recognize antigens on other enveloped viruses (14, 35). A potential explanation for the modest avidity-dependent increase is that only ≈10% of HIV spikes lie within the span of the 2 Fabs of an IgG (SI figure 2 in ref. 21), leaving most spikes available for only monovalent binding.

Second, the results suggested that the 4E10 epitope on gp41 is presented in a sterically constrained environment in contrast to the b12 epitope on gp120, which appeared to be fully accessible to potentially neutralizing reagents. Steric occlusion of the 4E10 epitope is consistent with the observation that a polymeric IgM version of 4E10 was significantly less potent than IgG 4E10 (36). The recent finding that 4E10 preferentially binds a fusion-intermediate conformation of gp41 (37) and that neutralization by IgG 4E10 is potentiated by the addition of a peptide that holds the trimer in a prehairpin intermediate state after attachment (38) provides additional context for interpretation of the steric occlusion effect, suggesting that the scFv and flexible scBvFv were better able to access a conformational state of the trimeric spike compared with the IgG, Fab, or diabody architectures. It is interesting to note, however, that occlusion effects were less evident for 1 of the tested strains, THRO4156.18. In this case, relative to scFv 4E10, IgG 4E10 and scBvFv 4E10 were 10-fold and 3.9-fold more potent, respectively

(Table S1), suggesting at least some ability for these bivalent architectures to mediate cross-linking. THRO4156.18 is also the only strain in which Fab 4E10 and scFv 4E10 were equally potent (Table S1). Together, these results suggest that strain THRO4156.18 might be suitable for vaccination efforts to raise 4E10-like antibodies.

Given the evidence for cross-linking by b12 reagents, we considered whether cross-linking occurred within the same spike trimer (intraspike) or between spike trimers (interspike). Analysis of a recent tomographic reconstruction of b12 Fabs bound to trimeric HIV-1 spikes on intact virions (20) suggests that intraspike cross-linking is not possible for IgG b12 or scBvFv b12 because the distance between 2 bound Fabs is greater than the span of either architecture (Fig. 1 and Fig. S3). The assumption that cross-linking is exclusively interspike allows us to address the potential for mobility of trimeric spikes on the viral surface, an issue that is relevant to the mechanisms of both antibody-mediated neutralization and fusion of the HIV-1 and host cell membranes. The low density of spikes on the surface of HIV-1 (20–23) would limit interspike cross-linking if the spikes were immobile or slow to diffuse relative to the kinetics of antibody binding. For both b12 and 4E10, we observed a greater avidity enhancement for the IgG architecture over the shorter scBvFv architecture (which span maximum distances of ≈15 and ≈11 nm, respectively; see Fig. 1), arguing in favor of a restriction to spike mobility.

The highly conserved MPER of gp41, which contains the 4E10 epitope, has long been considered an attractive target for vaccine design (39–47). Our observations of steric occlusion and inefficient cross-linking by IgG 4E10 for 7 of 8 primary isolates of clade B HIV-1 suggests that the IgG architecture is not optimal for bivalent recognition of its epitope, providing an explanation for the modest potency of 4E10 compared with other neutralizing antibodies (13). Combined with the low density of surface spikes on HIV-1, these limitations may serve as another mechanism by which HIV-1 limits neutralization by antibodies and represent an obstacle to vaccines that target the MPER. If so, IgG or scBvFv reagents with increased separation and/or flexibility between combining sites might represent a previously uncharacterized class of anti-HIV-1 reagents with increased neutralization potencies and therefore increased efficacy against HIV-1.

Materials and Methods

Affinity Determinations by Surface Plasmon Resonance. Proteins were produced (scFvs, scBvFvs, diabodies, Fabs, IgGs) or purchased (gp120, gp41) as described in *SI Text*. A Biacore 2000 biosensor system (Biacore International AB) was used to derive affinities of the b12 and 4E10 constructs for gp120 and gp41, respectively. In this assay, a protein (the "ligand") is covalently coupled to a gold–dextran layer, and association and dissociation phases for binding to injected protein (the "analyte") are measured in real time in resonance units (RU) (48, 49). The gp120 and gp41 proteins were immobilized by random primary amine coupling to a CM5 sensor chip as described in the Biacore manual. Monomeric gp120 was coupled at a density of 500 RU for experiments involving b12 constructs. gp41 was coupled at 300 RU for experiments involving scFv 4E10, scBvFv 4E10, and diabody 4E10 and at 150 RU for injections of IgG 4E10. A mock-coupled flow cell was used as a reference blank in all experiments. The surfaces were blocked with 3 5-min injections of 1 M ethanolamine (pH 8.0). After blocking, regeneration solutions of 60 mM

H_3PO_4 (for gp120) or 10 mM NaOH (for gp41) were repeatedly injected in short pulses until stable baselines were observed. Next, constant concentrations of the appropriate analytes followed by regeneration solution were repeatedly injected over both surfaces to verify reproducibility. For affinity measurements, a 2-fold dilution series of each analyte was injected over the flow cells at 100 μL/min at 25 °C in 10 mM Hepes buffer (pH 7.4), 150 mM NaCl, 3 mM EDTA, and 0.005% P-20 surfactant. Blank injections of just running buffer were used for double referencing (50). The chip surface was regenerated between analyte injections with 2 12-s injections.

Primary sensorgram data were preprocessed by using the Scrubber software package (Biologic Software; www.biologic.com.au). Kinetic constants were determined by simultaneously fitting the association and dissociation phases of all curves (4 or 5 injected concentrations per construct) to a 1:1 binding model by using ClampXP (51). For IgG 4E10, association data were collected at 4 concentrations, but dissociation data were collected for 2 h at only 1 concentration and fit separately because the dissociation rate was very slow. The 1:1 binding model describes a simple bimolecular interaction, yielding single association (k_{on}) and dissociation (k_{off}) values and a macroscopic (apparent) equilibrium dissociation constant (K_D), which includes density-dependent avidity effects that arise from the ability of bivalent constructs (the IgG, scBvFvs, and diabodies) to cross-link immobilized antigens. Because we wished to evaluate the effects of multivalent binding on the apparent affinities, we did not model the data for bivalent constructs with microscopic (stepwise) binding models because these models are defined in terms of monovalent binding events and would yield microscopic (i.e., intrinsic) affinities that do not include avidity effects. Errors for the K_D values were calculated with the formula $k_{off}/k_{on}*[(\delta_{on}/k_{on})^2 + (\delta_{off}/k_{off})^2]^{1/2}$, where δ_{on} and δ_{off} denote the asymptotic standard errors of the rate constants calculated in ClampXP.

Molecular Weight Determinations by Static Light Scattering. Static light-scattering experiments were performed at 25 °C by using a Superdex 75 10/30 gel filtration column (Amersham Biosciences) equipped with a Dawn Helios light scattering photometer and an Optilab rEX refractive index detector (Wyatt Technology). Protein samples (≈350 μg) were injected in TBS at a flow rate of 0.5 mL/min. Molecular weight values were calculated by using a dn/dc value of 0.185 mL/g. All data were analyzed with ASTRA software version 5.3.1.5 (Wyatt Technology).

Analysis of Neutralization Data. In vitro neutralization assays were conducted as described in the *SI Text* and previously (1, 32, 52). Molar 50% inhibitory concentration values (IC_{50}) were calculated by fitting the inhibition data to the equation $N = 100/[1+(IC_{50}/c)^H]$, where N is the percentage of neutralization, c is the concentration of the reagent being tested, and H is the Hill coefficient (KaleidaGraph v3.6, Synergy Software) (Fig. S2). For each antibody reagent, the mean IC_{50} value across 8 viral strains was calculated as an arithmetic mean by using the formula $\Sigma a_i/8$; $i = 1, 2, ..., 8$, where a_i refers to the IC_{50} value for viral strain i (Fig. 3), and as a geometric mean by using the formula $(\Pi a_i)^{(1/8)}$; $i = 1, 2, ..., 8$ (Fig. S4). The ratio of the IC_{50} value for a reagent compared with the IC_{50} value of another reagent was calculated as the ratio of the 2 means. Our conclusions did not differ using either type of calculation.

Structure Models. Models were created by using Swiss-PDB Viewer v3.9b2 (www. expasy.org/spdbv/) (53) and rendered in MacPymol (www.pymol.org/) (54).

1. Wei X, et al. (2003) Antibody neutralization and escape by HIV-1. *Nature* 422:307–312.
2. Kwong PD, et al. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393:648–659.
3. Labrijn AF, et al. (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.
4. Chen B, et al. (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* 433:834–841.
5. Kwong PD, et al. (2002) HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 420:678–682.
6. Burton DR, Stanfield RL, Wilson IA (2005) Antibody vs. HIV in a clash of evolutionary titans. *Proc Natl Acad Sci USA* 102:14943–14948.
7. Burton DR, et al. (1994) Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science* 266:1024–1027.
8. Barbas CF, III, et al. (1992) Recombinant human Fab fragments neutralize human type 1 immunodeficiency virus in vitro. *Proc Natl Acad Sci USA* 89:9339–9343.
9. Zwick MB, et al. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J Virol* 75:10892–10905.
10. Cardoso RM, et al. (2007) Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10. *J Mol Biol* 365:1533–1544.
11. Buchacher AR, et al. (1992) Human monoclonal antibodies against gp41 and gp120 as potential agent for passive immunization. *Vaccines* 92:191–195.

BIOCHEMISTRY

12. Muster T, et al. (1993) A conserved neutralizing epitope on gp41 of human immuno-deficiency virus type 1. *J Virol* 67:6642–6647.
13. Binley JM, et al. (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78:13232–13252.
14. Wu H, et al. (2005) Ultra-potent antibodies against respiratory syncytial virus: Effects of binding kinetics and binding valence on viral neutralization. *J Mol Biol* 350:126–144.
15. Icenogle J, et al. (1983) Neutralization of poliovirus by a monoclonal antibody: Kinetics and stoichiometry. *Virology* 127:412–425.
16. Cavacini LA, Emes CL, Power J, Duval M, Posner MR (1994) Effect of antibody valency on interaction with cell-surface expressed HIV-1 and viral neutralization. *J Immunol* 152:2538–2545.
17. Zhu Z, et al. (2006) Potent neutralization of Hendra and Nipah viruses by human monoclonal antibodies. *J Virol* 80:891–899.
18. Kausmally L, et al. (2004) Neutralizing human antibodies to varicella-zoster virus (VZV) derived from a VZV patient recombinant antibody library. *J Gen Virol* 85:3493–3500.
19. Janeway CA, Travers P, Walport M, Schlomchik MJ (2005) *Immunobiology* (Garland Science Publishing, New York), 6th Ed.
20. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109–113.
21. Zhu P, et al. (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441:847–852.
22. Sougrat R, et al. (2007) Electron tomography of the contact between T cells and SIV/HIV-1: Implications for viral entry. *PLoS Pathog* 3:e63.
23. Chertova E, et al. (2002) Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J Virol* 76:5315–5325.
24. Yamaguchi M, Danev R, Nishiyama K, Sugawara K, Nagayama K (2008) Zernike phase contrast electron microscopy of ice-embedded influenza A virus. *J Struct Biol* 162:271–276.
25. Bird RE, et al. (1988) Single-chain antigen-binding proteins. *Science* 242:423–426.
26. Huston JS, et al. (1988) Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli. Proc Natl Acad Sci USA* 85:5879–5883.
27. Mack M, Riethmuller G, Kufer P (1995) A small bispecific antibody construct expressed as a functional single-chain molecule with high tumor cell cytotoxicity. *Proc Natl Acad Sci USA* 92:7021–7025.
28. Bennett MJ, Eisenberg D (2004) The evolving role of 3D domain swapping in proteins. *Structure (London)* 12:1339–1341.
29. Holliger P, Prospero T, Winter G (1993) ''Diabodies'': Small bivalent and bispecific antibody fragments. *Proc Natl Acad Sci USA* 90:6444–6448.
30. Perisic O, Webb PA, Holliger P, Winter G, Williams RL (1994) Crystal structure of a diabody, a bivalent antibody fragment. *Structure (London)* 2:1217–1226.
31. Brunel FM, et al. (2006) Structure–function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. *J Virol* 80:1680–1687.
32. Li M, et al. (2005) Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* 79:10108–10125.
33. Sun ZY, et al. (2008) HIV-1 Broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane. *Immunity* 28:52–63.
34. Cardoso RM, et al. (2005) Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity* 22:163–173.
35. Schofield DJ, Stephenson JR, Dimmock NJ (1997) Variations in the neutralizing and haemagglutination-inhibiting activities of five influenza A virus-specific IgGs and their antibody fragments. *J Gen Virol* 78 (Pt 10):2431–2439.
36. Kunert R, Wolbank S, Stiegler G, Weik R, Katinger H (2004) Characterization of molecular features, antigen-binding, and in vitro properties of IgG and IgM variants of 4E10, an anti-HIV type 1 neutralizing monoclonal antibody. *AIDS Res Hum Retroviruses* 20:755–762.
37. Frey G, et al. (2008) A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. *Proc Natl Acad Sci USA* 105:3739–3744.
38. Gustchina E, Bewley CA, Clore GM (2008) Sequestering of the prehairpin intermediate of gp41 by peptide N36Mut(e,g) potentiates the human immunodeficiency virus type 1 neutralizing activity of monoclonal antibodies directed against the N-terminal helical repeat of gp41. *J Virol* 82:10032–10041.
39. Law M, Cardoso RM, Wilson IA, Burton DR (2007) Antigenic and immunogenic study of membrane-proximal external region-grafted gp120 antigens by a DNA prime-protein boost immunization strategy. *J Virol* 81:4272–4285.
40. Zwick MB (2005) The membrane-proximal external region of HIV-1 gp41: A vaccine target worth exploring. *AIDS* 19:1725–1737.
41. Muster T, et al. (1994) Cross-neutralizing activity against divergent human immuno-deficiency virus type 1 isolates induced by the gp41 sequence ELDKWAS. *J Virol* 68:4031–4034.
42. McGaughey GB, et al. (2003) HIV-1 vaccine development: Constrained peptide immu-nogens show improved binding to the anti-HIV-1 gp41 MAb. *Biochemistry* 42:3214–3223.
43. Liang X, et al. (1999) Epitope insertion into variable loops of HIV-1 gp120 as a potential means to improve immunogenicity of viral envelope protein. *Vaccine* 17:2862–2872.
44. Joyce JG, et al. (2002) Enhancement of alpha-helicity in the HIV-1 inhibitory peptide DP178 leads to an increased affinity for human monoclonal antibody 2F5 but does not elicit neutralizing responses in vitro. Implications for vaccine design. *J Biol Chem* 277:45811–45820.
45. Ho J, et al. (2005) Conformational constraints imposed on a pan-neutralizing HIV-1 antibody epitope result in increased antigenicity but not neutralizing response. *Vaccine* 23:1559–1573.
46. Eckhart L, et al. (1996) Immunogenic presentation of a conserved gp41 epitope of human immunodeficiency virus type 1 on recombinant surface antigen of hepatitis B virus. *J Gen Virol* 77 (Pt 9):2001–2008.
47. Phogat S, et al. (2008) Analysis of the human immunodeficiency virus type 1 gp41 membrane proximal external region arrayed on hepatitis B surface antigen particles. *Virology* 373:72–84.
48. Fagerstam LG, Frostell-Karlsson A, Karlsson R, Persson B, Ronnberg I (1992) Biospecific interaction analysis using surface plasmon resonance detection applied to kinetic, binding site and concentration analysis. *J Chromatogr* 597:397–410.
49. Malmqvist M (1993) Biospecific interaction analysis using biosensor technology. *Nature* 361:186–187.
50. Myszka DG (1999) Improving biosensor analysis. *J Mol Recognit* 12:279–284.
51. Myszka DG, Morton TA (1998) CLAMP: A biosensor kinetic data analysis program. *Trends Biochem Sci* 23:149–150.
52. Wei X, et al. (2002) Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* 46:1896–1905.
53. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18:2714–2723.
54. DeLano WL (2008) *The PyMOL Molecular Graphics System* (DeLano Scientific, Palo Alto, CA).
55. Saphire EO, et al. (2001) Crystal structure of a neutralizing human IGG against HIV-1: A template for vaccine design. *Science* 293:1155–1159.

Supporting Information

Klein et al. 10.1073/pnas.0811427106

SI Text

Cloning, Expression, and Protein Purification. Sequences encoding variable light and heavy (V_L and V_H) domains were amplified from genes encoding the b12 and 4E10 antibodies (gifts from D. R. Burton, Scripps Research Institute, La Jolla, CA). The V_L and V_H genes were fused using a linker encoding Gly$_4$Ser or (Gly$_4$Ser)$_3$ by bridge PCR to create 6 constructs (Fig. 1). Two were the monovalent b12 and 4E10 scFvs, which were constructed by using a (Gly$_4$Ser)$_3$ linker, and 4 were bivalent architectures derived from scFv genes. One of the bivalent architectures was a diabody, in which a shorter linker of only one Gly$_4$Ser repeat between the variable domains resulted in dimerization (diabodies b12 and 4E10) through 3D domain swapping (1–3). The other bivalent architecture was a single-chain bivalent Fv (scBvFv) consisting of a first scFv (V_L joined to V_H with a (Gly$_4$Ser)$_3$ linker) followed by (Gly$_4$Ser)$_3$ and a second scFv (V_H joined to V_L with a (Gly$_4$Ser)$_3$ linker).

For expression of monomeric scFvs, the b12 and 4E10 scFv genes were subcloned into the pET-22b(+) vector (EMD Biosciences) and transformed into *E. coli* strain BL21(DE3). Inclusion bodies containing unfolded scFv were solubilized and refolded as described previously (4). Briefly, bacterial cultures were grown to an OD_{600} of 0.9 at 37 °C, at which point IPTG was added to 1 mM, and the cultures were incubated for an additional 4 h. Inclusion bodies were isolated by 5 rounds of sonication and centrifugation, solubilized in 7 M guanidine with 10 mM reduced glutathione and 1 mM oxidized glutathione and refolded by rapid dilution at 4 °C in 0.1 M Tris·HCl (pH 8.0), 0.4 M arginine-HCl, 10 mM reduced glutathione, and 1 mM oxidized glutathione. After concentration, refolded scFv was further purified with Ni-NTA metal affinity Sepharose (Qiagen).

The gene encoding diabody 4E10 was subcloned into the bicistronic pAc-κ-Fc vector (PROGEN Biotechnik) for expression in baculovirus-infected insect cells. Genes encoding diabody b12 and the scBvFv constructs were subcloned into pAcGP67-A (BD PharMingen). Recombinant baculoviruses were generated by cotransfection of a transfer vector with linearized Baculogold viral DNA (BD PharMingen) and used to infect High Five cells (Invitrogen). Supernatants were then concentrated, buffer-exchanged with 50 mM Tris·HCl (pH 7.4), and 150 mM NaCl (TBS), and purified over Ni-NTA Sepharose.

IgGs were transiently expressed in mammalian cells. Full-length 4E10 heavy-chain (IgG1 subclass) and light-chain (κ) genes were subcloned separately into pcDNA3.1(−) (Invitrogen) and cotransfected into HEK 293T cells (American Type Culture Collection; ATCC) using Lipofectamine 2000 (Invitrogen). Constructs encoding the heavy chain and light chain of IgG b12 in the bicistronic vector pDR12 (a gift from D. R. Burton) were also expressed by transient transfection in HEK 293T cells. To produce Fab 4E10, a truncated 4E10 heavy-chain gene (terminated after residue Thr-252) was cotransfected into EBNA cells (ATCC) along with the 4E10 light-chain gene. Fab b12 was prepared by papain digestion of IgG b12 expressed in HEK 293F cells (Invitrogen) using 25-kDa linear PEI as a transfection reagent (Polysciences). Intact IgGs were purified by protein A chromatography (Pierce Biotechnology), and Fabs were purified using goat anti-human Fab polyclonal antibody (Sigma–Aldrich) cross-linked to Protein A beads.

All proteins were further purified by size-exclusion chromatography using Superdex 75 10/30, Superdex 75 16/60, or Superdex 200 26/60 columns (Amersham Biosciences) running in TBS. The final yields for bacterially expressed proteins were 2 and 0.5

mg/L for scFv b12 and scFv 4E10, respectively. The yields for scBvFv b12, diabody b12, scBvFv 4E10, and diabody 4E10 were 2.5, 1, 1.5, 0.7, mg/L of insect cell supernatant, respectively. The yields for IgG b12, IgG 4E10, and Fab 4E10, were 6.5, 0.8, and 2.6 mg/L, respectively.

Protein concentrations were determined by absorbance at 280 nm using extinction coefficients valid for denatured protein calculated from the secreted protein sequences using the online tool PlotParam (http://ca.expasy.org/tools/protparam.html) (5). Extinction coefficients were: scFv b12, 52370 $M^{-1}cm^{-1}$; scBvFv b12, 104740 $M^{-1}cm^{-1}$; diabody b12, 104740 $M^{-1}cm^{-1}$; Fab b12, 74425 $M^{-1}cm^{-1}$; IgG b12, 232542 $M^{-1}cm^{-1}$; scFv 4E10, 47900 $M^{-1}cm^{-1}$; scBvFv 4E10, 95800 $M^{-1}cm^{-1}$; diabody 4E10, 95800 $M^{-1}cm^{-1}$; Fab 4E10, 69330 $M^{-1}cm^{-1}$; IgG 4E10, 209480 $M^{-1}cm^{-1}$. No significant differences were observed when absorbance measurements were compared for scFv proteins diluted either in guanidine-HCl to 6.0 M or in 10 mM Hepes (pH 7.4) and 150 mM NaCl. All proteins were >95% pure and stable at concentrations of 1–2 mg/mL at 4 °C for at least several weeks as assessed by unchanged gel filtration profiles.

Purified recombinant full-length gp120 (clade B, strain HxBc2) expressed in Chinese hamster ovary cells and recombinant gp41 ectodomain (clade B, strain MN) expressed in bacteria were purchased from Immunodiagnostics. The recombinant gp41 is a 25-kDa fragment expressed as a fusion protein and purified by affinity chromatography and preparative electrophoresis (product specifications for Immunodiagnostics catalog no. 1091). To validate the antigenicity of the recombinant gp41, scFv and IgG versions of 2F5, a monoclonal antibody that binds an epitope near the 4E10-binding site on gp41 (6), were injected over the gp41 surface in the binding assay, resulting in a 10 nM affinity for the scFv, comparable with the 5.3 nM affinity observed when binding to the MPER peptide from gp41 and a low picomolar apparent affinity for IgG binding to gp41. No significant binding of scFv b12 to gp41 was observed at a concentration of 1 μM.

Strain Selection for in Vitro Neutralization Assays.

The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health: SVPB5, SVPB6, SVPB8, SVPB11, and SVPB12 (David Montefiori and Feng Gao); SVPB14, SVPB16, SVPB18 (B. H. Hahn and J. F. Salazar-Gonzalez); SVPB15 (B. H. Hahn and D. L. Kothe); SF162 (L. Stamatatos and C. Cheng-Mayer); SVPB17 (B. H. Hahn, X. Wei, and G. M. Shaw); pSG3$^{\Delta env}$ (John C. Kappes and Xiaoyun Wu); TZM-bl cells (John C. Kappes, Xiaoyun Wu, and Tranzyme Inc). HIV-1 pseudovirus particles from 10 pseudotyped primary virus strains [6535.3 (SVPB5), QH0692.42 (SVPB6), REJO4541.67 (SVPB16), RHPA4259.7 (SVPB14), SC422661.8 (SVPB8), SF162, THRO4156.18 (SVPB15), TRJO4551.58 (SVPB17), TRO.11 (SVPB12), and WITO4160.33 (SVPB18)] were prepared, and in vitro neutralization assays were performed as described previously (7–9).

To be able to generalize our results across an entire clade, we examined the sensitivities of various clades to both b12 and 4E10 using previously published results (10) to select appropriate strains. 4E10 neutralizes most/all strains regardless of clade, so choosing appropriate clade(s) for our comparison depended on b12, which is less broadly neutralizing. It was observed that clades A, B, C, D, AE, and BG contained at least 1 strain that was sensitive to IgG b12 (10). To be useful in our experiments

in which potentially weakly neutralizing reagents (e.g., monovalent constructs) would be tested, we estimated that a strain would need to have an $IC_{90} < 50$ μg/mL for IgG b12 in order for it to be possible for us to accurately derive an IC_{50} for a monovalent reagent. Clade B was the only clade in which a majority of strains met this criterion (Clade A: 2 of 12 strains; clade B: 19 of 29; clade C: 5 of 12; clade D: 3 of 11; clade AE: 1 of 10; clade BG: 1 of 1). We therefore chose to confine our comparative studies to clade B.

In Vitro Neutralization Assay Methods. Briefly, TZM-bl cells (8, 11, 12), a HeLa cell line expressing CD4, the HIV-1 coreceptors CCR5 and CXCR4, and Tat-responsive firefly luciferase, were infected by pseudotyped viruses, and single rounds of infection were detected as luminescence from luciferase. Each antibody reagent was tested for inhibition of infection in triplicate by preincubating 5,000 infectious viral units per well with a 3-fold dilution series of the potential inhibitor for 1 h at 37 °C. Ten thousand TZM-bl cells were then added to each well. After 48 h at 37 °C, the cells were lysed in the presence of Bright-Glo (Promega), and relative luminescence was recorded by using a Victor3 luminometer (PerkinElmer). Percent neutralization was calculated as $[1 - [(\delta - \gamma)/(\phi - \gamma)]] \times 100$, where δ is the relative luminescence observed for each sample well, γ is the background luminescence observed for cells without virus or antibody reagent, and ϕ is the maximum luminescence observed for cells with virus only.

Neutralization curves were plotted as the percentage of neutralization (y axis) versus concentration of potential inhibitor (x axis). Each data point on a neutralization curve is the mean of a triplicate measurement $\pm$ SEM. IC_{50} values were calculated as described (13). Briefly, neutralization curves were fit to the equation $N = 100/[1 + (IC_{50}/c)^{H}]$ where N is percent neutralization and c is the concentration of the reagent being tested, which constrains the maximum and minimum of each curve to 100% and 0% neutralization, respectively, and H represents the Hill coefficient (KaleidaGraph v3.6; Synergy Software). Fitting the data with a Hill coefficient constrained to a value of 1 does not change the results. Errors reported for the IC_{50} values in Table 1 were the asymptotic standard error calculated from nonlinear regression analyses and therefore represent the goodness-of-fit. An IC_{50} value could not be determined for diabody b12 neutralization of strain WITO4160.33 because the protein was not stable at the concentration necessary to achieve neutralization above $\approx$50%.

To evaluate the reproducibility of the IC_{50} values, we used either scFv b12 or IgG b12 as internal controls. When data from these multiple replicates were available, we reported the IC_{50} value in Table 1 as the average, and a representative curve is shown in Fig. S2. The error reported in Table 1 was then calculated as the product of the average IC_{50} value and the square root of the average of the sum of squares of the fractional errors from each of the replicates. The results from these multiple independent replicates are summarized in Table S2 and a comparison of our IC_{50} values with published IC_{50} values for IgG b12 and IgG 4E10 are presented in Table S3. These comparisons demonstrate that our neutralization curves yielded reproducible IC_{50} values that are in agreement with those published by others.

1. Bennett MJ, Eisenberg D (2004) The evolving role of 3D domain swapping in proteins. *Structure (London)* 12:1339–1341.
2. Perisic O, Webb PA, Holliger P, Winter G, Williams RL (1994) Crystal structure of a diabody, a bivalent antibody fragment. *Structure (London)* 2:1217–1226.
3. Holliger P, Prospero T, Winter G (1993) "Diabodies": Small bivalent and bispecific antibody fragments. *Proc Natl Acad Sci USA* 90:6444–6448.
4. Steinle A, et al. (2001) Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics* 53:279–287.
5. Gasteiger E, et al. (2005) Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*, ed Walker JM (Humana Press, Totowa, NJ), pp 571–607.
6. Muster T, et al. (1993) A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* 67:6642–6647.
7. Li M, et al. (2005) Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* 79:10108–10125.
8. Wei X, et al. (2002) Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob Agents Chemother* 46:1896–1905.
9. Wei X, et al. (2003) Antibody neutralization and escape by HIV-1. *Nature* 422:307–312.
10. Binley JM, et al. (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78:13232–13252.
11. Derdeyn CA, et al. (2000) Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *J Virol* 74:8358–8367.
12. Platt EJ, Wehrly K, Kuhmann SE, Chesebro B, Kabat D (1998) Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol* 72:2855–2864.
13. Gustchina E, Bewley CA, Clore GM (2008) Sequestering of the prehairpin intermediate of gp41 by peptide N36Mut(e,g) potentiates the human immunodeficiency virus type 1 neutralizing activity of monoclonal antibodies directed against the N-terminal helical repeat of gp41. *J Virol* 82:10032–10041.
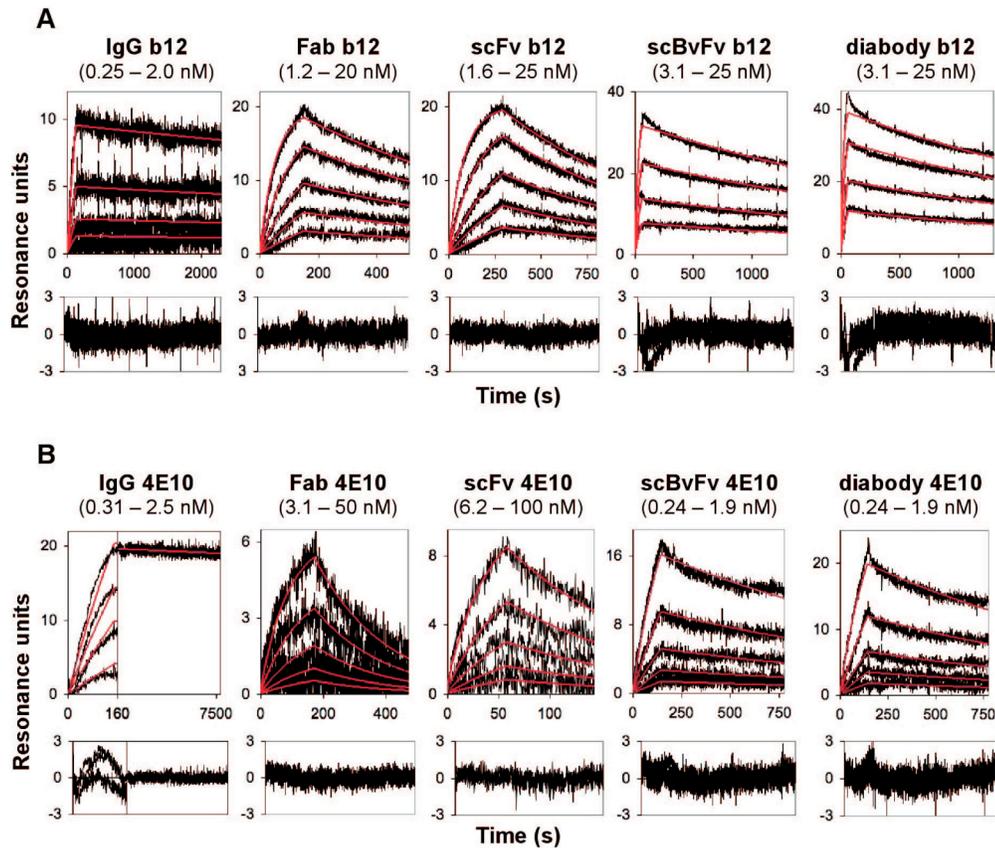
Fig. S1. Surface plasmon resonance sensorgrams for binding to immobilized antigens. Analytes were injected with the range of concentrations listed above the sensorgrams. (*A*) Sensorgrams for 2-fold dilution series of the indicated analyte binding to immobilized monomeric gp120. (*B*) Sensorgrams for 2-fold dilution series of the indicated analyte binding to immobilized gp41 ectodomain. All sensorgrams were fit with a 1-to-1 binding model. For IgG 4E10 binding to gp41, the off-rate (a concentration-independent parameter) was fit separately from the on-rate. Residual plots are shown beneath each set of fitted sensorgrams.

42
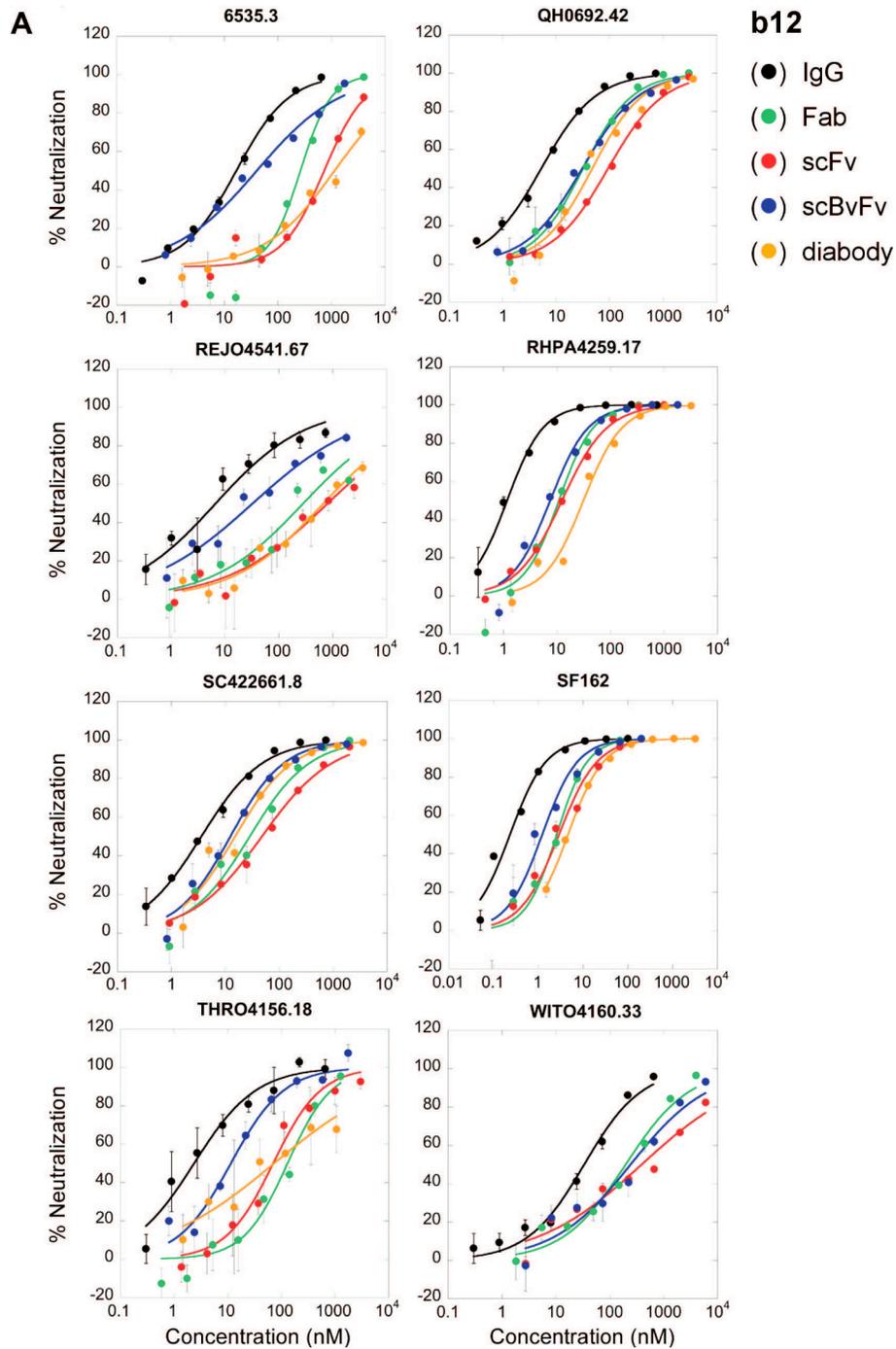
Fig. S2. In vitro pseudovirus neutralization curves. (*A*) b12 constructs.
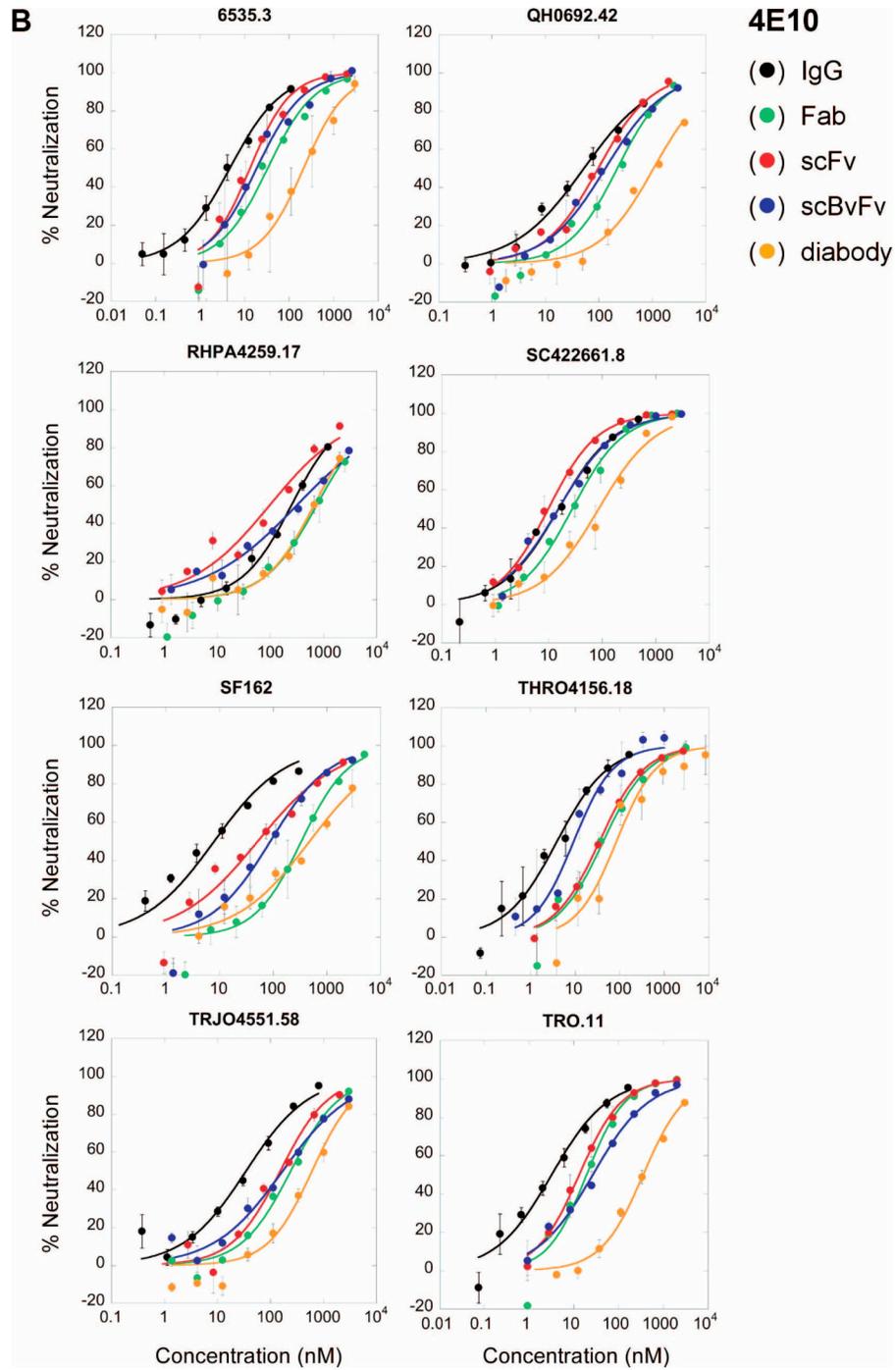
Fig. S2B. 4E10 constructs.

Top view
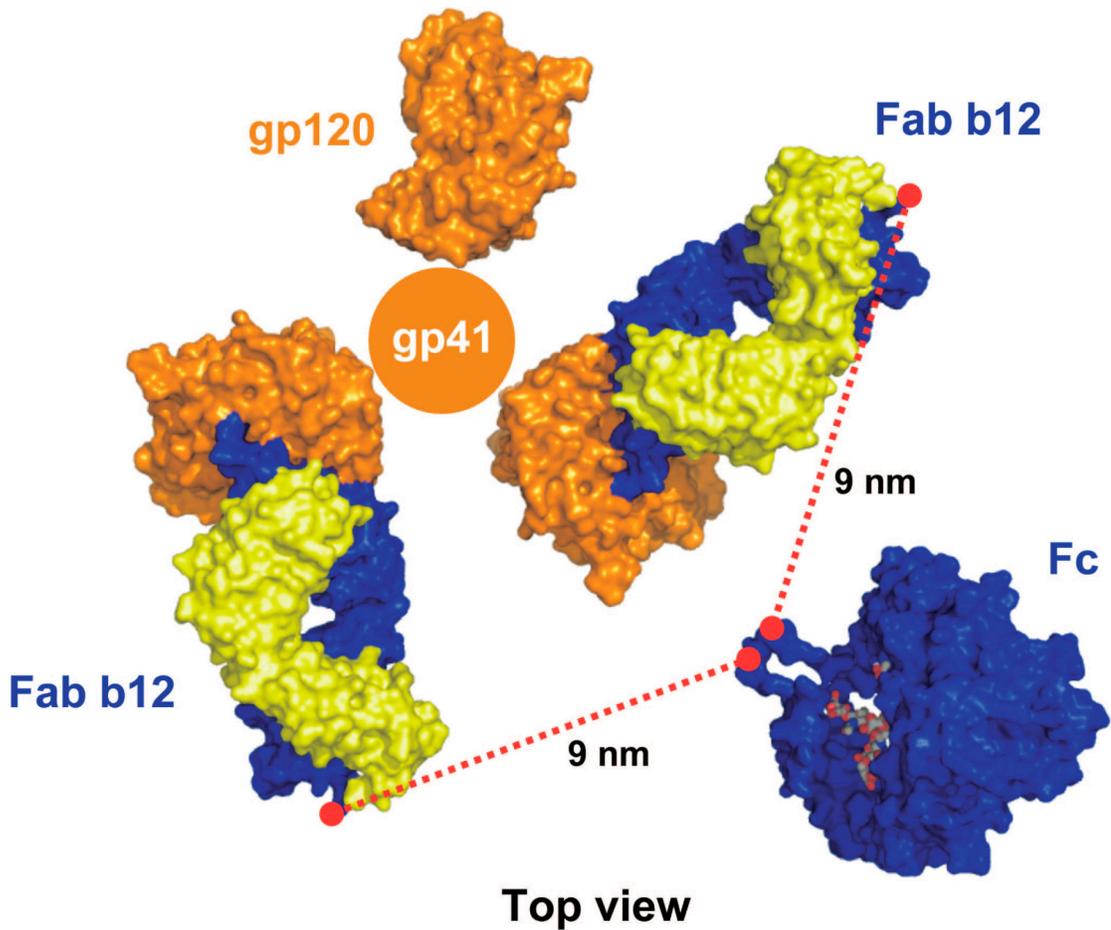
Fig. S3. Modeling of the structural requirements for IgG b12 to achieve intraspike cross-linking. Using the coordinates of the HIV gp120 trimer in its b12-bound state (orange) derived from tomographic reconstructions of intact HIV trimers (PDB ID code 3DNL) (1), we aligned the coordinates for Fab b12 (heavy chain in blue and light chain in yellow) using the crystal structure of a Fab b12/gp120 complex (PDB ID code 2NY7) (2). The Fc domain from the crystal structure of IgG b12 (PDB ID code 1HZH) (3) was then placed between and approximately equidistant from the 2 Fabs. The distance between the N-terminal Cys residue of each Fc chain (red dot) and the C-terminal Cys residue of each Fab heavy chain (red dot) is ≈9 nm, which is ≈8 nm longer than a typical IgG hinge (e.g., see IgG in Fig. 1).

1. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109–113.
2. Zhou T, et al. (2007) Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* 445(7129):732–737.
3. Saphire EO, et al. (2001) Crystal structure of a neutralizing human IGG against HIV-1: A template for vaccine design. *Science* 293(5532):1155–1159.
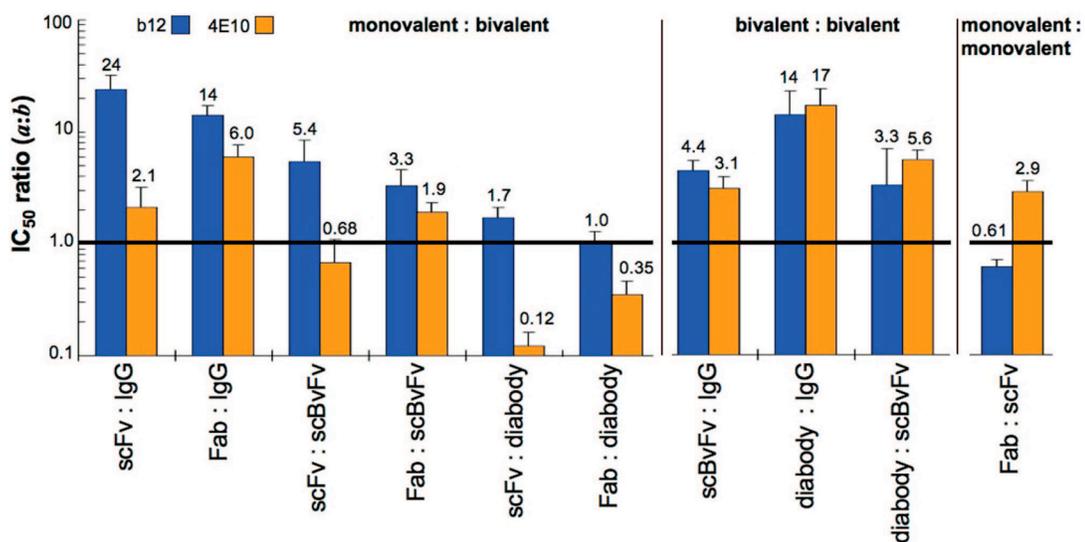
Fig. S4. Bar graph of ratios of average molar IC$_{50}$ values (geometric means) for b12 constructs (blue) and 4E10 constructs (orange). Reagent pairs with an average ratio of 1.0 (black line) are equal in average potencies. Ratios >1.0 indicate that reagent b is more potent than reagent a. Ratios <1.0 indicate that reagent a is more potent than reagent b. Error bars represent the standard errors calculated from the variability in strain-specific ratios for each pair of reagents.

Table S1. Strain-specific IC$_{50}$ neutralization ratios

				IC$_{50}$ ratio by virus strain				
b12	6535.3	QH0692.42	RHPA4259.7	SC422661.8	SF162	THRO4156.18	REJO4541.67	WITO4160.33
Monovalent/bivalent								
scFv/IgG	40 ± 16	20 ± 2	12 ± 3	15 ± 3	19 ± 7	26 ± 10	81 ± 30	18 ± 6
Fab/IgG	14 ± 5	14 ± 2	11 ± 3	7.4 ± 2.1	10 ± 4	30 ± 12	29 ± 12	14 ± 3
scFv/scBvFv	19 ± 7	3.5 ± 0.5	1.6 ± 0.1	4.8 ± 1.0	4.1 ± 2.0	9.2 ± 2.3	25 ± 9	1.6 ± 0.6
Fab/scBvFv	6.3 ± 2.2	2.5 ± 0.4	1.5 ± 0.3	2.3 ± 0.7	2.2 ± 1.1	11 ± 3	8.8 ± 3.6	1.3 ± 0.4
scFv/diabody	0.69 ± 0.28	2.3 ± 0.5	0.63 ± 0.08	3.8 ± 1.1	1.1 ± 0.4	1.8 ± 0.9	1.2 ± 0.4	n.d.
Fab/diabody	0.24 ± 0.08	1.6 ± 0.4	0.58 ± 0.12	1.9 ± 0.7	0.57 ± 0.22	2.1 ± 1.1	0.43 ± 0.17	n.d.
Bivalent/bivalent								
scBvFv/IgG	2.2 ± 0.5	5.5 ± 0.7	7.5 ± 1.5	3.2 ± 0.7	4.6 ± 1.7	2.8 ± 1.0	3.3 ± 1	11 ± 3.5
diabody/IgG	58 ± 14	8.6 ± 1.9	19 ± 4.3	3.9 ± 1.2	18 ± 3	14 ± 8	68 ± 20	n.d.
diabody/scBvFv	27 ± 6	1.5 ± 0.4	2.5 ± 0.3	1.3 ± 0.4	3.8 ± 1.3	5.2 ± 2.6	21 ± 6	n.d.
Monovalent/monovalent								
Fab/scFv	0.34 ± 0.16	0.69 ± 0.11	0.92 ± 0.18	0.49 ± 0.14	0.53 ± 0.27	1.2 ± 0.3	0.36 ± 0.17	0.82 ± 0.3
4E10	6535.3	QH0692.42	RHPA4259.7	SC422661.8	SF162	THRO4156.18	TRJO4551.58	TRO.11
Monovalent/bivalent								
scFv/IgG	2.8 ± 0.9	2.6 ± 0.4	0.40 ± 0.11	0.67 ± 0.15	2.9 ± 1.7	10 ± 3	2.6 ± 0.9	2.5 ± 0.5
Fab/IgG	6.8 ± 2.4	5.8 ± 1.5	3.0 ± 1.0	1.9 ± 0.4	15 ± 9	12 ± 4	6.9 ± 1.6	5.9 ± 1.9
scFv/scBvFv	0.74 ± 0.20	0.75 ± 0.20	0.31 ± 0.07	0.67 ± 0.11	0.62 ± 0.31	3.9 ± 0.5	0.54 ± 0.19	0.30 ± 0.05
Fab/ scBvFv	1.8 ± 0.5	1.7 ± 0.5	2.4 ± 0.7	1.9 ± 0.3	3.2 ± 1.7	4.7 ± 1.2	1.4 ± 0.3	0.70 ± 0.20
scFv/diabody	0.064 ± 0.029	0.089 ± 0.020	0.15 ± 0.04	0.19 ± 0.04	0.12 ± 0.07	0.42 ± 0.12	0.15 ± 0.05	0.040 ± 0.011
Fab/diabody	0.15 ± 0.05	0.2 ± 0.06	1.1 ± 0.3	0.52 ± 0.15	0.60 ± 0.27	0.50 ± 0.26	0.39 ± 0.09	0.10 ± 0.03
Bivalent/ bivalent								
scBvFv/IgG	3.8 ± 1.2	3.4 ± 0.9	1.3 ± 0.3	1.0 ± 0.2	4.7 ± 2.4	2.6 ± 0.7	4.9 ± 1.3	8.4 ± 1.8
diabody/IgG	44 ± 21	29 ± 6	2.7 ± 0.7	3.6 ± 0.9	25 ± 15	25 ± 9	17 ± 5	63 ± 20
diabody/scBvFv	12 ± 5	8.5 ± 2.5	2.1 ± 0.4	3.6 ± 0.8	5.3 ± 2.6	9.3 ± 2.8	3.6 ± 1.0	7.4 ± 2.0
Monovalent/monovalent								
Fab/scFv	2.4 ± 0.8	2.2 ± 0.6	7.6 ± 2.5	2.8 ± 0.4	5.2 ± 3.1	1.2 ± 0.3	2.6 ± 0.8	2.4 ± 0.7

The average IC$_{50}$ ratios presented in Fig. 3 of the main text were calculated as the ratio of their respective arithmetic mean IC$_{50}$ values as opposed to an arithmetic mean of the individual ratios presented here. n.d., not done. See discussion in *SI Text*.

Table S2. IgG b12 and scFv b12 were used as internal controls to examine the reproducibility of independently determined IC_{50} values

Isolate	IC_{50} replicates, nM				Average
	1	2	3	4	
IgG b12					
6535.3	7.4 ± 1.4	17 ± 2	19 ± 2	34 ± 6	19 ± 3
QH0692.42	4.6 ± 0.9	5.3 ± 0.6	6.6 ± 0.3	6.7 ± 0.6	5.8 ± 0.7
RHPA4259.7	0.61 ± 0.15	1.2 ± 0.1	1.2 ± 0.1	—	1.0 ± 0.2
SC422661.8	2.1 ± 0.3	2.8 ± 0.5	3.7 ± 0.5	6.9 ± 1.3	3.9 ± 0.6
SF162	0.23 ± 0.04	0.31 ± 0.08	—	—	0.27 ± 0.06
THRO4156.18	2.3 ± 0.5	5.2 ± 1.2	5.4 ± 2.2	—	4.3 ± 1.4
scFv b12					
6535.3	410 ± 160	610 ± 340	740 ± 180	1400 ± 400	790 ± 310
QH0692.42	86 ± 17	95 ± 9	120 ± 20	130 ± 20	110 ± 20
RHPA4259.7	12 ± 1	12 ± 1	—	—	12 ± 1
SC422661.8	36 ± 8	51 ± 10	89 ± 14	—	59 ± 11
SF162	3.0 ± 1.6	4.4 ± 0.6	7.6 ± 1.4	—	5.0 ± 1.7
THRO4156.18	71 ± 10	150 ± 30	—	—	110 ± 20

Table S3. Comparison of IC$_{50}$ values for IgG b12 and IgG 4E10 to previously published results

Isolate	Ours (μg/mL)	Li, et al. (μg/mL)*	Fold difference
IgG b12			
6535.3	2.5	1.4	1.8
QH0692.42	0.87	0.3	2.9
RHPA4259.7	0.15	0.1	1.5
SC422661.8	0.59	0.2	3.0
SF162	0.04	0.01 (0.03[†])	4.1 (1.4[†])
THRO4156.18	0.66	0.5	1.3
REJO4541.67	1.5	0.7	2.1
WITO4160.33	3.3	3.1	0.9
IgG 4E10			
6535.3	0.72	0.2	3.6
QH0692.42	5.5	1.4	3.9
RHPA4259.7	38	6.9	5.5
SC422661.8	1.5	0.9	1.7
SF162	2.7	0.3 (4.0[†])	9.0 (0.7[†])
THRO4156.18	0.51	0.3	1.7
TRJO4551.58	5.4	4.5	1.2
TRO.11	0.48	0.3	1.6

*With the exception of SF162, the HIV isolates used in our study were initially characterized in Li, et al. (1).
[†]Our IC$_{50}$ values for this strain more closely match those reported by Binley, et al. (2), the laboratory in which the original characterization of this strain was conducted.

1. Li M, et al. (2005) Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol* 79:10108–10125.
2. Binley JM, et al. (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78:13232–13252.

CHAPTER **3**

Assessing the impact of HIV spike arrangement

on antibody avidity

This chapter presents a hypothesis called the "island effect" that may explain the poor efficacy of existing broadly neutralizing anti-HIV antibodies and the ongoing failure to develop a vaccine capable of eliciting any substantial neutralizing antibody response, and data are presented on the development of antibody hinge extension technologies that could eventually lead to novel antibody architectures capable of overcoming this effect. This work was completed with the assistance of Priyanthi Gnanapragasm, Rachel Galimidi, Chris Foglesong, and Maria Suzuki.

Introduction

It is estimated that since 1990, there have been more than 50 million HIV infections (1). Yet despite a well-developed global network of treatment and monitoring facilities, to date, the only documented case in which an established infection was apparently cleared was after complete ablation of the hematopoietic stem cells (HSCs) of an HIV$^+$ patient undergoing treatment for leukemia followed by replacement with donor HSCs bearing a homozygous deletion of CCR5 (2). Even after HAART-mediated long-term suppression of virus replication to undetectable levels, viral loads eventually rebound after treatment interruption (3), suggesting that it is not possible for a human immune system that has been pre-primed with an infection to enforce HAART-equivalent suppression.

Rapid mutation is a hallmark of HIV infection (4) and is logically invoked as an important factor fueling successful immune evasion that, coupled with the structural features of the envelope trimer, helps explain the paucity of evidence to suggest that the antibody response significantly contributes to the relatively low viral set point observed after the acute viremia phase of infection. Nevertheless, as discussed in previous chapters, conserved epitopes within the envelope trimer have been identified along with broadly neutralizing antibodies (bNAbs) that recognize them. One might predict that passively immunizing an HIV patient by injecting one or more of these bNAbs could lead to indefinite suppression of an infection and/or prevent infection in viral challenge models. Indeed, clinical trials using a cocktail of three bNAbs (2G12, 4E10, and 2F5) demonstrated a partial ability to suppress viral rebound among patients who underwent interruptions in HAART (5). However, the relatively brief delays in viral rebound that

were observed required serum concentrations for the bNAbs totaling ~2 mg/mL. Considering that the volume of blood in a human adult is ~5 liters and that the serum half-lives of most antibodies are at best several weeks, one would need to inject unrealistic quantities of antibodies to achieve even the partial protection seen in this study, where researchers had to inject 1 to 1.3 grams of each antibody weekly (6). The short-lived delay in viral rebound was particularly surprising given the fact that the serum concentrations of each of the individual antibodies ranged up to two orders of magnitude higher than the *in vitro* IC_{90} values for the patients' autologous viruses.

Another study that highlighted the poor performance of bNAbs examined the concentrations required to protect rhesus macaques from infection using chimeric SIV/HIV ($SHIV_{SF162P3}$) derived from strain SF162, which is exceptionally sensitive to the antibody b12 (7), exhibiting an *in vitro* IC_{50} of less than 0.1 µg/mL. Complete protection for these animals from infection by the chimeric virus was observed for dosages at 25 mg/kg, which yielded serum concentrations of approximately 0.7 mg/mL. The concentration of IgG in humans is ~10 mg/mL (8); thus, ~7 % of the total serum IgG would need to be composed of b12 in order to achieve protection against a variant of HIV that was shown to be 50 times more sensitive than the median to b12 among a panel of 19 strains (Fig. 1). These results suggest that if one were to select a typical strain of HIV for the infection challenge, it would probably be necessary to inject enough b12 to more than double the total concentration of IgG in order to achieve protection.
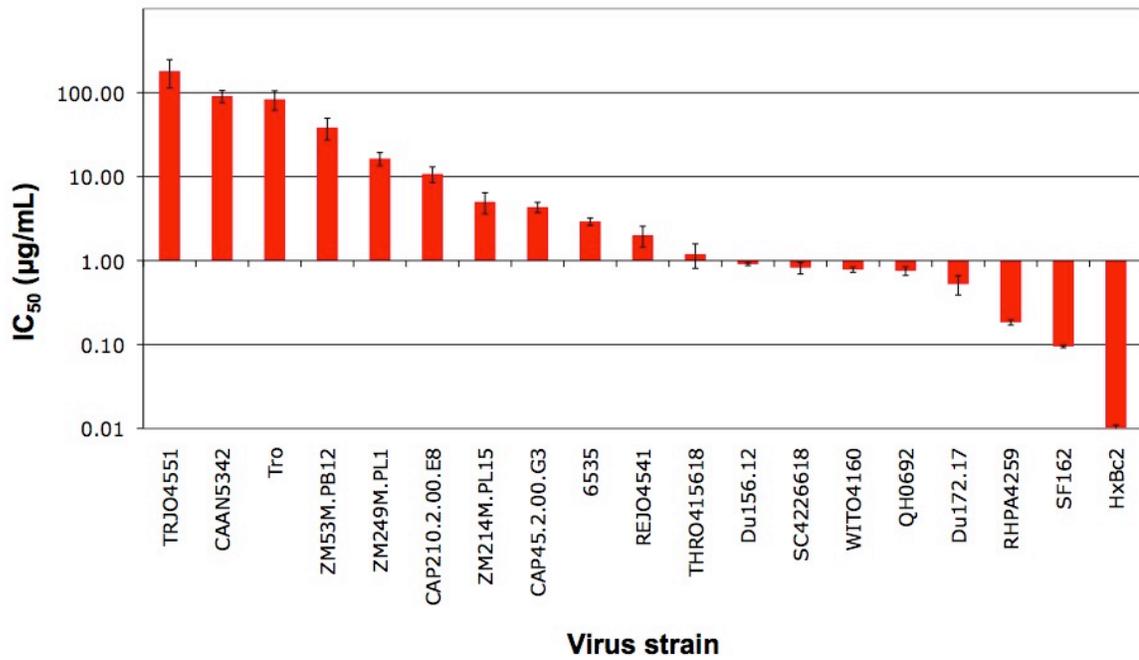
Figure 1. IC$_{50}$ values for the neutralization of various strains of HIV from clades B and C using b12 including SF162, second bar from the right (Klein JS, Gnanapragasm PNP, Björkman PJ, unpublished).

The biophysics of antibody-antigen interactions

The mediocre antibody performance observed in both the human and NHP studies against strains of virus that are unusually sensitive to them suggest that an alternative explanation to rapid mutation may be undermining the capacity of these antibodies to neutralize HIV. To explore this possibility, it is worthwhile to review the structural and biophysical aspects of antibody-antigen interactions.

Intrinsic (*i.e.*, monovalent) antibody-antigen binding mechanisms can often be described using a simple one-step binding model. Consider the interaction between an anti-gp120 Fab and monomeric gp120. In this interaction the Fab, *A*, binds to gp120, *B*, to form a reversible complex, *AB,* which can be described with single rate constants for the forward and reverse reactions, k_1 and k_{-1}:

$$A + B \underset{k_{-1}}{\overset{k1}{\rightleftharpoons}} AB \tag{1}$$

The rate equations that describe the chemical equations can be easily derived from the law of mass action, which states that the rate of any reaction is proportional to the product of the concentration(s) of the reactant(s). Therefore, the instantaneous rates for the formation of *B* and *AB* can be written as:

$$-\frac{d[B]}{dt} = \frac{d[AB]}{dt} = k_1[A][B] - k_{-1}[AB] \tag{2}$$

where [*A*], [*B*], and [*AB*] denote the molar concentrations of the Fab, gp120, and the complex, respectively. This model can be used to derive the intrinsic equilibrium dissociation constant, k_D, which can be calculated from either the ratio of the concentrations of reactants to products when the system is at equilibrium (Eq. 3), or as the ratio of the reverse reaction rate to the forward reaction rate (Eq. 4):

$$k_D = \frac{[A][B]}{[AB]} \tag{3}$$

$$k_D = \frac{k_{-1}}{k_1} \tag{4}$$

In the more complex scenario of binding between one intact bivalent antibody and one stationary envelope spike where intra-spike cross-linking cannot occur (as discussed in Chapter 2), an intrinsic binding model does not account for the effect of having multiple pathways to the formation of the complex, *AB*. Here, binding can occur between the spike and either Fab. Therefore, two equivalent pathways lead to the formation of *AB* and one pathway leads to the formation of *B*, yielding the potential for a two-fold increase in the apparent (*i.e.,* observed) affinity, K_D:

$$K_D = \frac{k_{-1}}{k_1 + k_1} = \frac{k_D}{2} \tag{6}$$

The upper limits of diffusion that have been measured for both Fabs and IgGs, which have been shown to be reduced at distances of up to ~100 nm from planar surfaces and spherical membranes relative to their rates in bulk solution, is between 10^5 and 10^6 $M^{-1}s^{-1}$ (9). Consistent with these findings and the results presented in Chapter 2, values of k_1 for high-affinity antibody-antigen interactions are generally reported to be within the bounds of this upper limit (10-14) despite attempts to develop mutants with enhanced association rates (15). Consequently, in order to achieve mid- to low-picomolar intrinsic affinities, it is necessary for k_{-1} to be $\leq 10^{-4}$ s^{-1} even with the statistical increase in the apparent forward reaction rate that may be observed with the presence of two relatively unconstrained Fab arms.

While antibodies with intrinsic affinities in the nanomolar range are reasonably common, there are few reports of antibodies with intrinsic affinities in the sub-nanomolar

range. However, when both Fab arms are able to engage two binding sites that are fixed in space relative to each other (such as may be observed on the surface of a SPR sensor chip or a pathogen), the resulting increase in apparent affinity can lead to nearly irreversible binding even though the intrinsic affinity may be significantly higher than 1 nM (*e.g.*, see Fig. 2*C* in Chapter 2). This property, called avidity, arises from the relatively low probability that both Fab arms will both occupy a dissociated state for a long enough period of time to allow complete dissociation of the complex. Avidity can be modeled as a two-step binding interaction between a bivalent analyte (the antibody, or *A*) and an immobilized ligand (the antigen, or *B*), which can form the complexes *AB* and *AB₂*:

$$A + B \underset{k_{-1}}{\overset{k1}{\rightleftharpoons}} AB \tag{7}$$

$$AB + B \underset{k_{-2}}{\overset{k2}{\rightleftharpoons}} AB_2 \tag{8}$$

Thus, a system of non-linear differential equations describe the instantaneous rates for the formation of *B, AB,* and *AB₂*:

$$\frac{d[B]}{dt} = -k_1[A][B] + k_{-1}[AB] - k_2[AB][B] + k_{-2}[AB_2] \tag{9}$$

$$\frac{d[AB]}{dt} = k_1[A][B] - k_{-1}[AB] - k_2[AB][B] + k_{-2}[AB_2] \tag{10}$$

$$\frac{d[AB_2]}{dt} = k_2[AB][B] - k_{-2}[AB_2] \tag{11}$$

Modeling the rate of formation of *B* from *AB* using Eq. 2 for monovalent binding (Fab) and from *AB₂* using Eqs. 9-11 for bivalent binding (IgG) by numerical analysis demonstrates the slow apparent dissociation rate that arises from perfectly efficient cross-linking (Fig. 2 and Table 1).
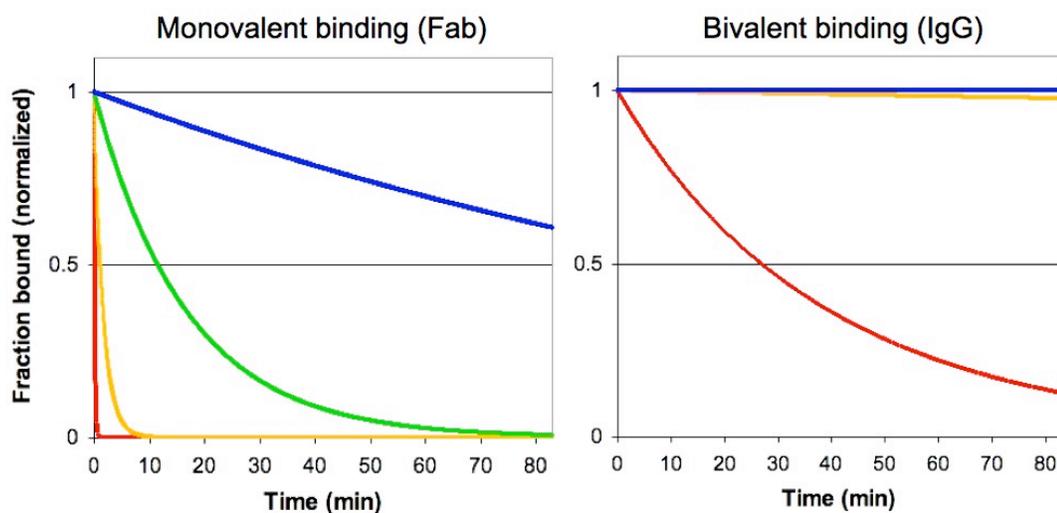
Figure 2. Modeling the dissociation to *B* from *AB* (Fab) and *AB₂* (IgG). In both panels, dissociation phases are calculated for 85 minutes. Intrinsic association rates were fixed at $10^5 \ M^{-1}s^{-1}$ and models were generated using a range of values for intrinsic dissociation rates of $10^{-1} \ s^{-1}$ to $10^{-4} \ s^{-1}$, yielding intrinsic equilibrium dissociation constants of 1 μM (•), 100 nM (•), 10 nM (•), and 1 nM (•). Simulated curves of bivalent binding with intrinsic affinities of 100 nM, 10 nM, and 1 nM were nearly identical such that the curves appear indistinguishable.

Table 1. Half-life values ($t_{1/2}$) of antibody-antigen complexes calculated from models presented in Fig. 2. Values for dissociation involving a monovalent interaction (Fab) were calculated using the equation $t_{1/2} = \ln(2)/(60k_{-1})$. Half-life values for dissociation involving a bivalent interaction (IgG) were derived by inspection or by approximation using the equation for first-order exponential decay.

		$t_{1/2}$ (min)	
k_{-1} (s⁻¹)	k_D (nM)	Fab	IgG
10^{-1}	1000 nM	0.12	27
10^{-2}	100 nM	1.2	~3,000
10^{-3}	10 nM	12	~200,000
10^{-4}	1 nM	120	> 200,000

The importance of avidity to virus neutralization was elegantly demonstrated in a study of poliovirus, a non-enveloped icosahedral virus containing 30 two-fold symmetric axes. The authors were able to demonstrate that antibody binding saturates near the expected value of 30 bivalent antibodies per virus particle and that cleaving the antibodies such that the Fab arms became unlinked led to a substantial increase in the molar concentration required to inhibit infection (16). Thus, the bivalent nature of IgG antibodies facilitates binding to and neutralization of viruses.

Evidence for avidity enhancement in the neutralization of enveloped viruses

In the absence of bivalent binding (*i.e.*, cross-linking) or confounding factors such as an increase in the potency of an IgG relative to one of its component Fabs arising simply from its larger size, the maximum difference in neutralization potency between these architectures should be two-fold for the reasons stated previously. Thus, the direct comparison of neutralization potencies for Fabs and IgGs against different viruses can be used as a semi-quantitative assay for cross-linking efficiency.

If cross-linking is necessary for antibody-mediated neutralization at concentrations that are therapeutically relevant, then it should be possible to isolate antibodies that exhibit efficient cross-linking from individuals that have successfully cleared pathogens or been effectively vaccinated against them. Vaccines have been developed for hepatitis B, measles, and influenza type A, all of which are enveloped viruses like HIV. In addition, passively administered antibody therapies exist for both hepatitis B and respiratory syncytial virus (RSV) – another enveloped virus. In stark contrast to what has been observed for HIV, comparisons of IgG antibodies and their Fab fragments reactive to RSV

and influenza type A have demonstrated molar differences in neutralization that are in excess of 1000-fold (Fig. 3). The demonstration that some of these viruses are permissive to highly efficient cross-linking is indicative of the possibility that such viruses share structural features absent from HIV.

Given that avidity primarily derives its effect by limiting the rate of dissociation, a prediction for the difference in potency between an IgG and a Fab as a function of k_{-1} would be that as the value for k_{-1} decreases so should the difference in potencies between the two architectures. For example, one group of researchers created a library of mutants of the antibody Palivizumab, which neutralizes RSV and is used as a prophylaxis (15). Consistent with this prediction, the authors found that as they decreased the intrinsic k_{-1}, thereby increasing the neutralization potency of the Fab, the neutralization potency of the IgG remained relatively unchanged (15) (Fig. 4). In this sense, *one may think of avidity as acting as a buffer to de-optimization of the antibody-antigen interface* (*e.g.,* a pathogen that evades antibody binding by mutation), further underscoring the role that avidity could be playing in antibody-mediated neutralization.

In asking the question of how the topology of spikes might differ between HIV and other enveloped viruses in such a way as to make them more permissive to efficient cross-linking by antibodies, it is helpful to examine the available data from studies by electron microscopy. Cryo-preserved virus samples can be imaged at low-nanometer resolutions such that individual proteins may be resolved on the surfaces of the viruses and reconstructed in three-dimensional space by tomography. Tomographic reconstructions of whole HIV particles have only recently become available (17-19), but they are consistent with older biochemical data that showed a relatively low abundance of spikes on the viral surface (20).
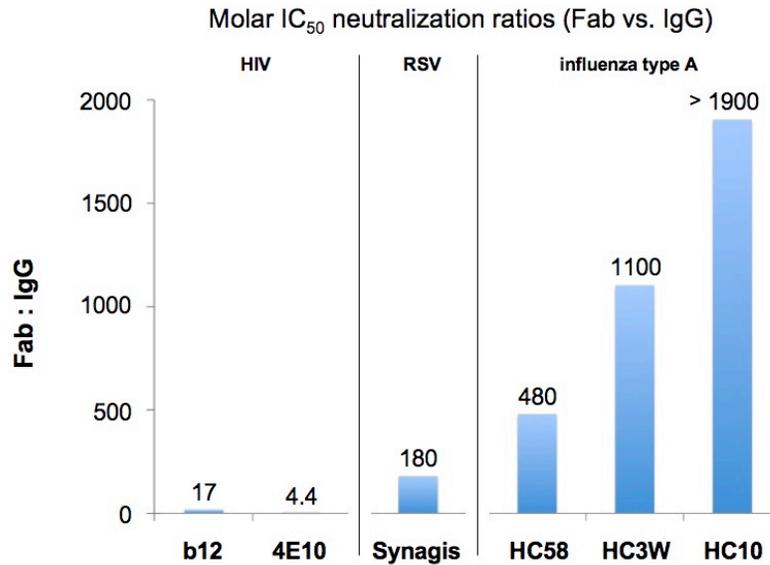
Figure 3. Comparison of IC_{50} neutralization values for antibodies as Fabs and intact IgGs. Molar ratios were calculated as the molar IC_{50} value for the Fab divided by the molar IC_{50} value for the IgG (RSV, respiratory syncytial virus (15); IA, influenza type A (21)).
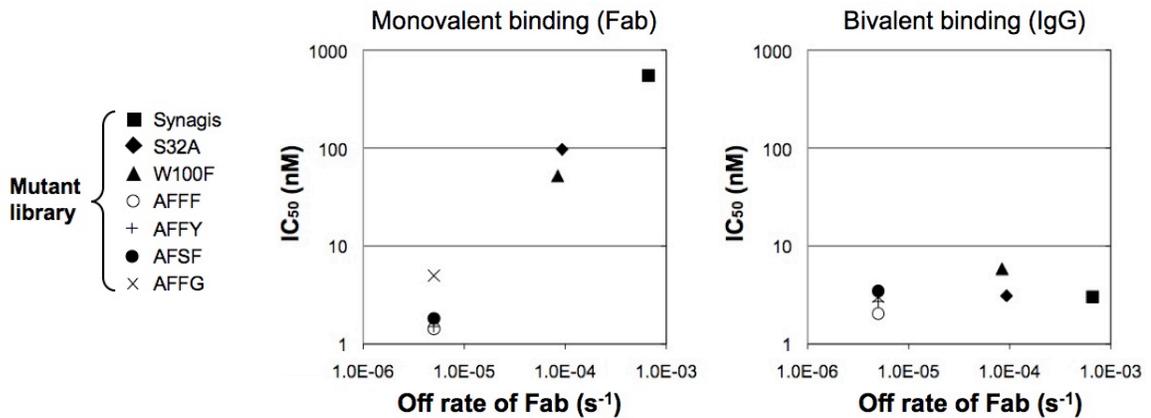
Figure 4. Comparison of neutralization potencies for Palivizumab and affinity-improved mutants with slower intrinsic dissociation constants (reported in (15)). IC_{50} values were determined for the antibodies as Fabs and as intact IgGs.

The Island Effect

Enveloped viruses contain a cell-derived membrane in which viral antigens are acquired during budding from the host cell. An examination of electron micrographs of enveloped viruses for which antibody-mediated neutralization is known to be critical to the control and/or elimination of infection (22-24) reveals a high density of viral spikes on the surfaces of these viruses (Fig. 5). For example, influenza type A virus contains ~450 spikes, which are spaced at intervals $\leq$ 10 nm (Fig. 5*B*). Similarly, the spikes on measles and hepatitis B virus are closely spaced (Fig. 5*A* and 5*C*). However, biochemical studies and three-dimensional electron microscopic reconstructions revealed that HIV appears to have only 14 ± 7 spikes per virus particle (17, 18, 20, 25) (Fig. 5*D*), probably arising tandem endocytosis motifs that limit the concentration of spikes on the cell surface and the labile nature of the non-covalent gp120-gp41 complex (26, 27).

An analysis of nearest neighbor distances between individual spikes on HIV particles revealed that most of the spikes are separated by distances that far exceed the 12-15 nm reach of the two Fab arms of an IgG (Fig. 5*E*) (17). Given evidence of interactions between gp41 and the viral matrix protein of HIV (28, 29) as well as restricted spike mobility and the inability to cross-link epitopes within a spike (see Chapter 2), these large inter-spike separations may represent an insurmountable barrier to efficient cross-linking by naturally produced anti-HIV antibodies. This model predicts that the only exceptions to the poor efficacy of a monoclonal anti-HIV antibody should be for the condition where the intrinsic affinity of the Fab for a particular strain of HIV is strong enough to be comparable to the avidity-enhanced affinity that would be observed under conditions of efficient cross-linking.
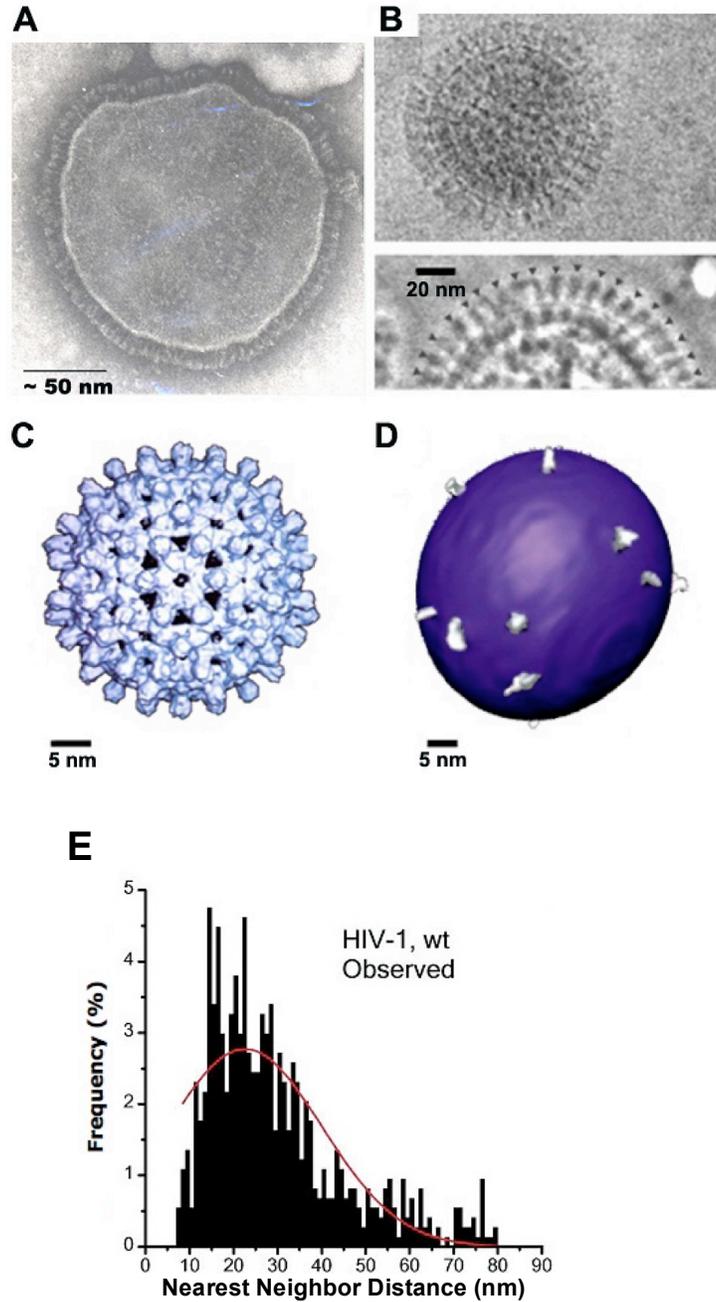
Figure 5. Available structures of enveloped viruses showing spike densities derived by electron microscopy. (*A*) Measles virus (adapted from (30)). (*B*) Influenza type A. Black triangles denote individual spikes in the lower panel (adapted from (31)). (*C*) Hepatitis B virus (adapted from (32)). (*D*) HIV (adapted from (17)). (*E*) Bar graph of the distribution of nearest neighbor distances between HIV envelope spikes (adapted from (17)).

An exception to the island effect: anti-carbohydrate antibodies

Approximately 25 N-linked glycosylation sites are conserved on gp120 (33), although the locations of the sites may vary between strains, such that approximately half of its molecular weight is composed of carbohydrate (34). This feature likely presents anti-carbohydrate antibodies an opportunity to cross-link antigen within a single viral spike. Although antibodies that recognize viral carbohydrates are rare because they are made by host enzymes and therefore expected to be non-immunogenic, one potent bNAb against HIV, 2G12, does recognize a constellation of viral carbohydrates within a single gp120 subunit (35). This unexpected antigenicity is probably related to the mannose content arising from inefficient processing of sterically occluded high-mannose N-linked glycans (36).

Carbohydrate recognition by 2G12 is accomplished using an unusual domain-swapped structure in which the two Fabs create a single antigen recognition region with two rigidly arranged binding sites separated by ~3.5 nm (37). We recently isolated and characterized a naturally occurring dimeric form of 2G12 containing four Fabs instead of two, and the dimeric form exhibited ~50-fold increased neutralization potency across a range of clade B HIV strains compared to monomeric 2G12 (Appendix F). Although it has been shown that gp120 can easily mutate to change its carbohydrate topology (33), which may explain why 2G12 does not show broad cross-reactivity across known strains of HIV (38), the apparent ability of 2G12 to efficiently cross-link antigen makes anti-carbohydrate antibodies with specificities similar to 2G12 a potential source of potent (though perhaps clade-specific) antibodies.

63

Circumventing the island effect

If the island effect hypothesis is correct, it predicts that attempts to raise NAbs against HIV by injection of purified antigens or attenuated or killed viruses are predestined to fail because too few of the elicited antibodies will bind with high enough intrinsic affinities to the virus to be of practical use. Even if such antibodies were elicited, exposure to strains of HIV whose relevant epitopes do not adequately match those of the strain used for vaccination, a likely scenario, are likely to substantially limit their prophylactic value because they will lack the buffering effect of avidity against natural variation. However, novel antibody architectures may provide a path to negating the island effect, as it may be possible to construct a bivalent antibody in which both of its Fabs are capable of simultaneously binding to epitopes within a single virus spike, eliminating the challenge of accommodating the wide range of nearest neighbor distances between spikes.

The following results summarize ongoing efforts to develop such architectures. Thus far, two different strategies have been examined as candidates for a new hinge extension technology: random coil extension (RCE), and protein fusion extension (PFE). The RCEs were encoded using the $(Gly_4Ser)_x$ sequence (39). Candidates for the PFEs were derived using well-characterized proteins for which high resolution X-ray crystal structures exist, showing that they adopt extended conformations – such candidate proteins included CD4 (40), invasin and a truncation thereof (41), a sequence composed of immunoglobulin domains 168-170 from titin (42), and a tandem repeat of two fibronectin type III domains from neogenin (Yang F, Björkman PJ, unpublished results) (Fig. 6).

A

CD4
titin$_{168-170}$
invasin
invasin$_{1-3}$
neogenin (FNIII$_{5-6}$)$_2$

12 nm
12 nm
13 nm
18 nm
19 nm

B

CD4

titin$_{168-170}$
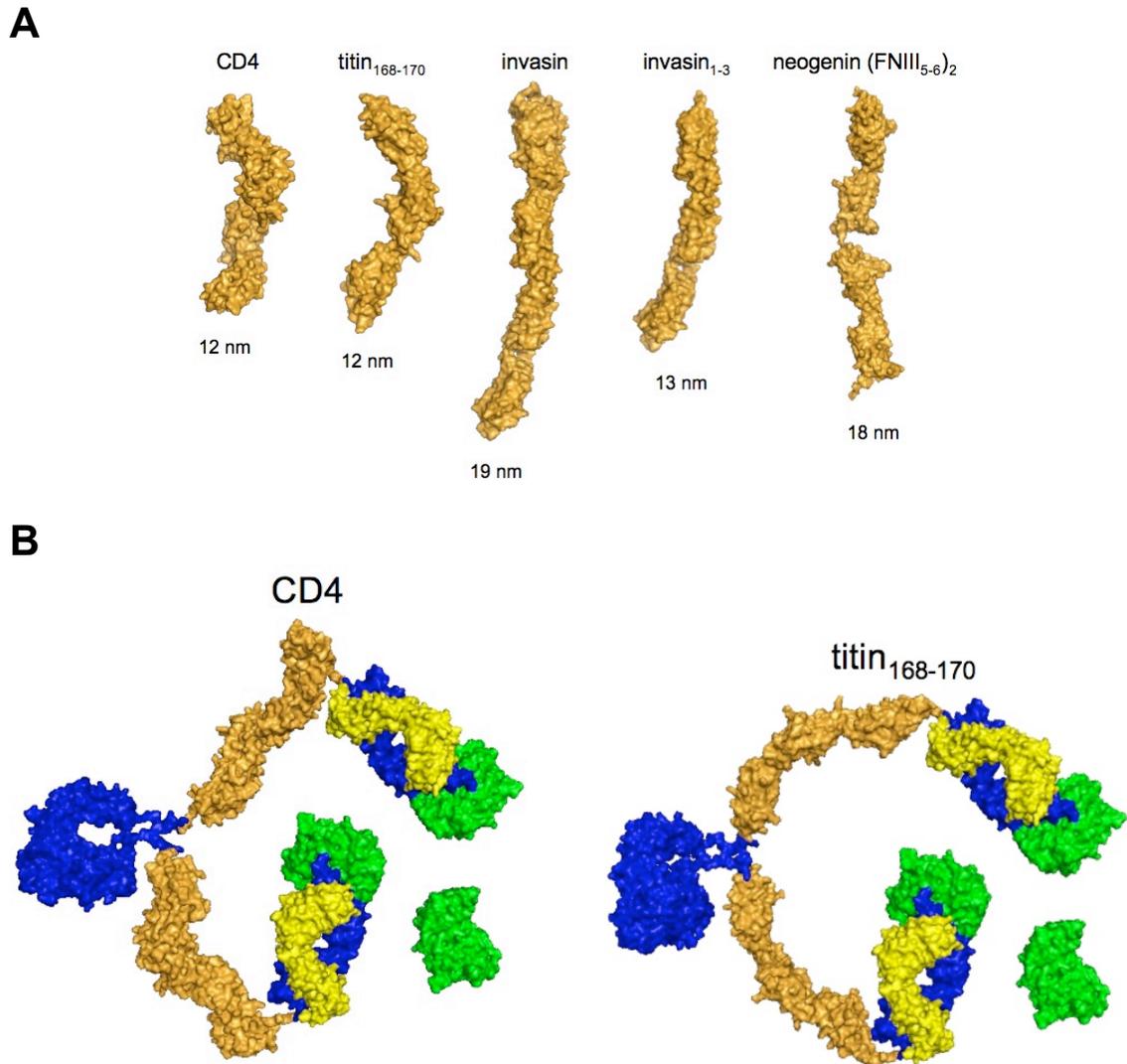
Figure 6. Protein fusion linker candidates and models. (*A*) Candidates for protein fusion linkers differing in length and flexibility (*i.e.*, the hinge domain of CD4 or the two identical neogenin fragments linked by a GlySer sequence versus the relatively rigid structure of invasin and a truncated form, invasin$_{1-3}$). (*B*) These top view models are examples of protein fusion extensions (PFE) and are based on the trimeric spike model presented in Chapter 2 using the coordinates from high-resolution X-ray crystal structures of gp120 docked to available tomographic reconstructions of intact spikes on cryo-preserved virus particles (gp120, green; PFE, gold; Fab, blue and yellow; Fc, blue).

Results

Characterization of IgGs with RCEs

IgG forms of b12, 2G12, 4E10, and 2F5 were expressed with RCEs 1, 3, 5, and 7 repeats of a Gly_4Ser sequence inserted into the hinge region separating the CH_1 domain from the CH_2 domain in an IgG1 architecture N-terminal to the naturally occurring DKTHT hinge sequence. All proteins were purified by Protein A affinity chromatography followed by SEC. When analyzed by reduced SDS PAGE, slight increases in the molecular weights of the heavy chains with increasing hinge length were observed (Fig. 7*B*).

Protein expression decreased as a function of increased linker length for constructs with 1, 3, and 5 Gly_4Ser repeats, and little to no expression was observed for constructs with 7 repeats for some IgGs (Table 2). This sharp decline in expression was observed for all antibodies tested, suggesting that RCEs of this length introduce a general limit to the folding of IgG sequences or might inhibit dimerization either between the heavy chains or between the heavy chains and light chains. In addition, reduced amounts of 2G12 dimer were recovered from the RCE expressions with more than one Gly_4Ser insertion, indicating that extended hinge sequences prevent the intermolecular domain swapping observed for the wild type architecture of this antibody (Appendix F). Although the sequences were codon optimized, more recent advances in codon optimization algorithms may provide an opportunity to offset the losses in protein yield (43, 44).
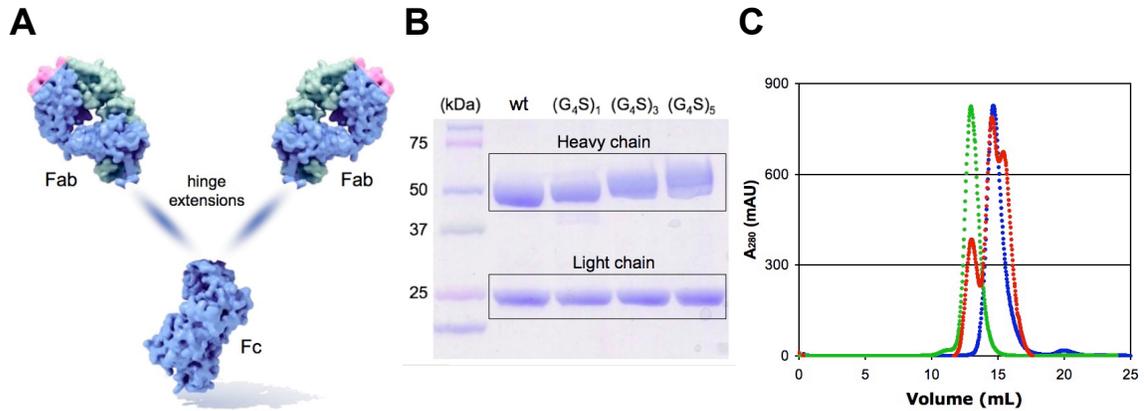
Figure 7. (*A*) Schematic of the structure of a human IgG antibody of subclass 1 with RCEs. Pink coloring denotes the locations of antigen recognition sites (adapted from U.S. National Library of Medicine). (*B*) Reduced SDS PAGE results for IgG b12 with Gly$_4$Ser RCEs of increasing length. (*C*) Gel filtration profile of Fab 2G12 (blue), IgG 2G12 (G$_4$S)$_1$ after purification (green) and after one month of storage in TBS at 4 °C (red).

Table 2. Protein yields for 1 L expressions of each of the RCE constructs.

Antibody	wt	(Gly$_4$Ser)$_1$	(Gly$_4$Ser)$_3$	(Gly$_4$Ser)$_5$	(Gly$_4$Ser)$_7$
		Total yield (mg/L) and percentage of wild type yield			
b12	6	6 (100%)	6 (100%)	5 (83%)	0.5 (8%)
2G12	20	15 (75%)	5 (25%)	4 (20%)	2 (10%)
4E10	7	4 (57%)	3 (43%)	2 (28%)	0
2F5	4	2 (50%)	1.5 (38%)	1 (25%)	0

It was also observed that precipitation or degradation into Fab and Fc fragments occurred following storage for periods of > 1 month at 4 °C in TBS (Fig. 7C), yet neither the IgGs nor the scFvs described in Chapter 2 exhibited degradation after storage for similar lengths of time. Interestingly, these effects were not observed for several of the $(Gly_4Ser)_5$ RCEs that had been stored as Protein A eluates (a citrate buffer neutralized with Tris) and purified by SEC in TBS only after long-term storage. Thus, it might be possible to inhibit the apparent instability of RCEs with alternative buffer conditions.

In theory, the ability of an IgG to bind with avidity will be reduced as the hinge length increases because it reduces the effective concentration of the Fab arms (45). To examine whether the RCE antibodies exhibited this predicted effect, wild type IgG, RCEs with the $(Gly_4Ser)_5$ extension (the longest one that expressed efficiently) and monomeric Fab forms were injected over a CM5 surface displaying 500 RU of either monomeric gp120 or gp41. Half-life times ($t_{1/2}$) were then calculated from their apparent dissociation rate constants by approximation with a model for first-order exponential decay (Fig. 8 and Table 3). No significant differences in $t_{1/2}$ values were observed between the wild type and RCE antibodies, demonstrating that the ability to bind with avidity remained intact. As expected, the $t_{1/2}$ values for all of the Fabs were significantly faster with the exception of 2G12. The similar rate of dissociation for the 2G12 Fab as compared to the bivalent constructs is likely due to the domain swapped structure remaining intact.

To examine whether RCEs exhibited an avidity enhancement during neutralization, which would be reflected as a decrease in the IC_{50} values relative to the wild type IgG, the potency of each of the antibodies were compared in an *in vitro*
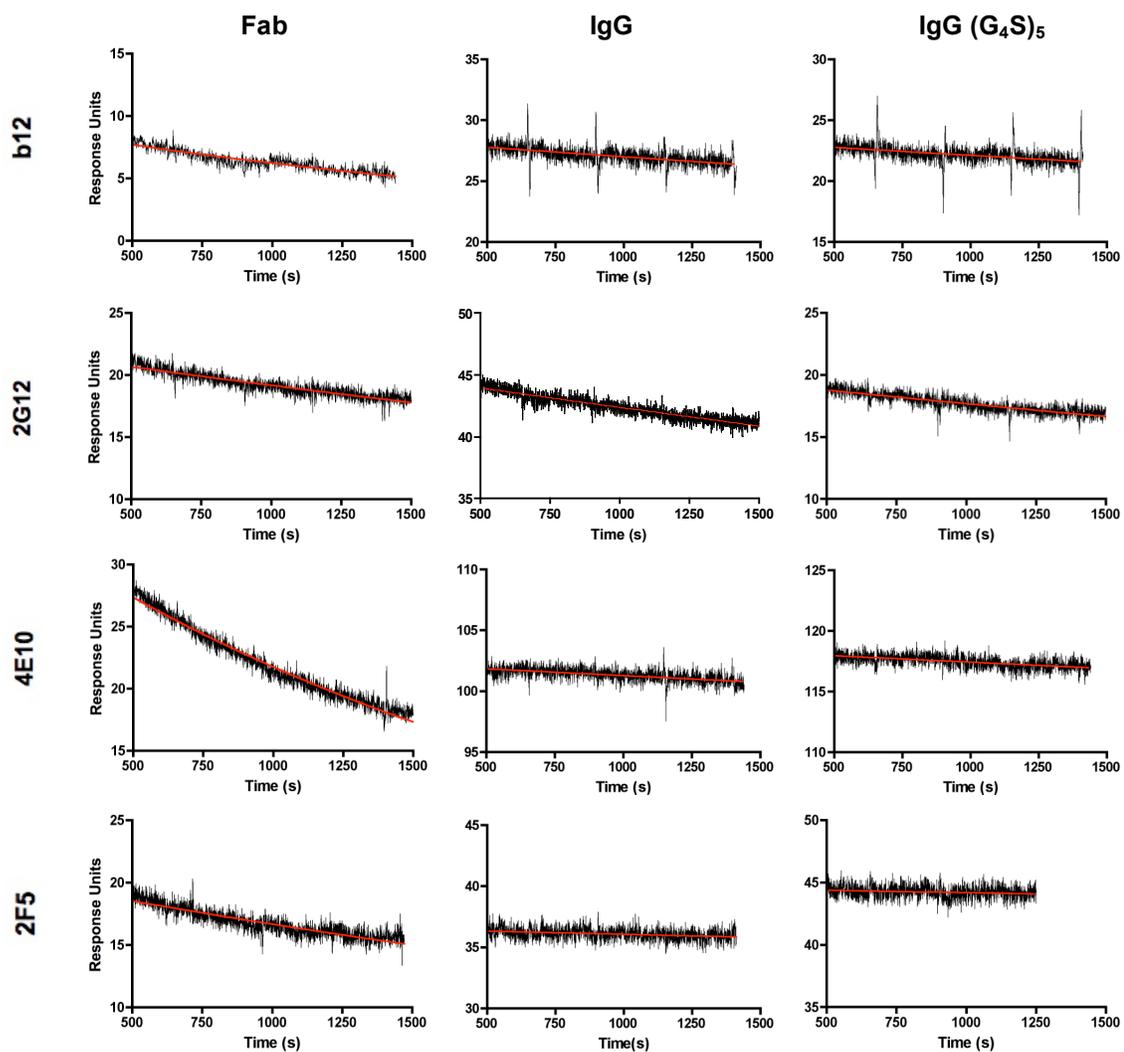
Figure 8. SPR dissociation curves for Fab, IgG, and IgG (G₄S)₅ constructs bound to immobilized monomeric gp120 (b12 and 2G12) or immobilized gp41 (4E10 and 2F5).

Table 3. Half-life times ($t_{1/2}$) calculated from dissociation curves in Fig. 8.

Antibody	Fab	IgG wt	IgG (G$_4$S)$_5$
		$t_{1/2}$ (min)	
b12	28	200	200
2G12	77	160	96
4E10	25	1000	1300
2F5	55	800	1300
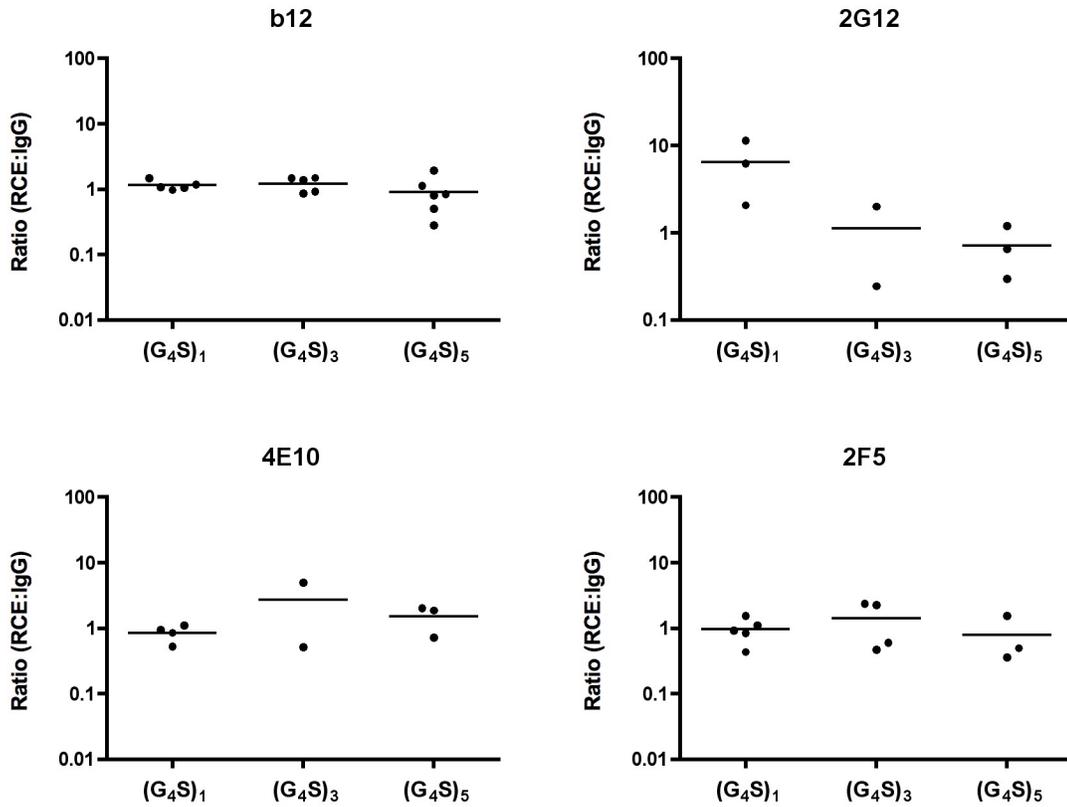
Figure 9. Results of the *in vitro* neutralization assay comparing wild type IgG constructs with RCEs. Values were calculated as the ratio of the molar IC_{50} determined for the RCE divided by the molar IC_{50} determined for the parental IgG. The IgG 2G12 $(Gly_4Ser)_3$ construct was non-neutralizing for strain QH0692.42.

neutralization assay. Each of the antibody constructs was tested against virus strains from clade B (Fig. 9).

When comparing the molar IC_{50} values for the RCEs with the values determined for the wild type IgGs, nearly all were within a few fold of each other. One significant exception was observed for 2G12 versus strain QH0692.42. Whereas wild type IgG 2G12 neutralized this strain with an IC_{50} of 59 nM, the 2G12 RCE with a single Gly_4Ser repeat was 6-fold less potent (IC_{50} value of 360 nM) and the RCE with three Gly_4Ser repeats was completely non-neutralizing. It was also observed that Fab 2G12 is also non-neutralizing against this strain.

Characterization of IgGs with PFEs

Expression tests were completed for all of the protein fusion extensions (PFEs). Each construct was purified by protein A affinity chromatography followed by SEC. While some of the 4E10 and 2G12 PFEs and one for 2F5 yielded between 0.5 and 1 mg/L (Figs. 10 and 11), none of the b12 constructs yielded measurable quantities of purified protein. The b12 constructs were of most interest because it was the Fab b12 structure bound to the trimeric spike that was used for the rational design of the PFEs (Fig. 6*B*). This indicates that for future studies it may be necessary to start with simpler designs that express well and then build on those. Thus, a set of small single domain candidates to serve as protein fusions are currently in development, which, provided they express well, will be used in multi-domain PFEs, each separated by short flexible linkers.

The PFEs that did express were evaluated for their abilities to bind immobilized antigen by SPR and for their abilities to neutralize HIV *in vitro* (Fig. 12 and Table 4). For

both 4E10 and 2F5, the PFEs that did express at levels sufficient to test (invasin and invasin$_{1-3}$ for 4E10 and invasin$_{1-3}$ for 2F5), sensograms for binding to immobilized gp41 were comparable to their respective wild type IgGs (Table 4). However, whereas the kinetics for dissociation of IgG 2G12 with an invasin$_{1-3}$ PFE were observed to be comparable relative to wild type IgG, the rate of dissociation for the 2G12 neogenin PFE was observed to be significantly slower ($t_{1/2}$ of 1200 min versus 160 min, Tables 3 and 4). Interestingly, dimeric IgG 2G12, which was also examined by SPR (Fig. 12 and Table 3), showed a slightly faster rate of dissociation than its monomeric form ($t_{1/2}$ of 65 min). Thus, in spite of its ~50-fold greater potency in neutralization (Appendix F), there appears to be no evidence of an avidity enhancement for dimeric IgG 2G12 when binding monomeric gp120 immobilized on a CM5 surface, suggesting that the enhanced potency of dimeric 2G12 derives from an ability to bind glycans on multiple gp120 monomers.

When examined for their respective abilities to neutralize HIV (strain 6535.3), the PFEs were considerably less potent than their respective wild type IgGs (Fig. 13). These results support the findings presented in Chapter 2 of steric occlusion for the 4E10 epitope in that the short linkage between the Fab and the PFE (-GGSGGSA-) may further inhibit access to its epitope. However, these data represent the first evidence of a similar effect for 2F5, whose epitope is immediately N-terminal to the 4E10 epitope, indicating that the MPER as a whole may be sterically occluded.
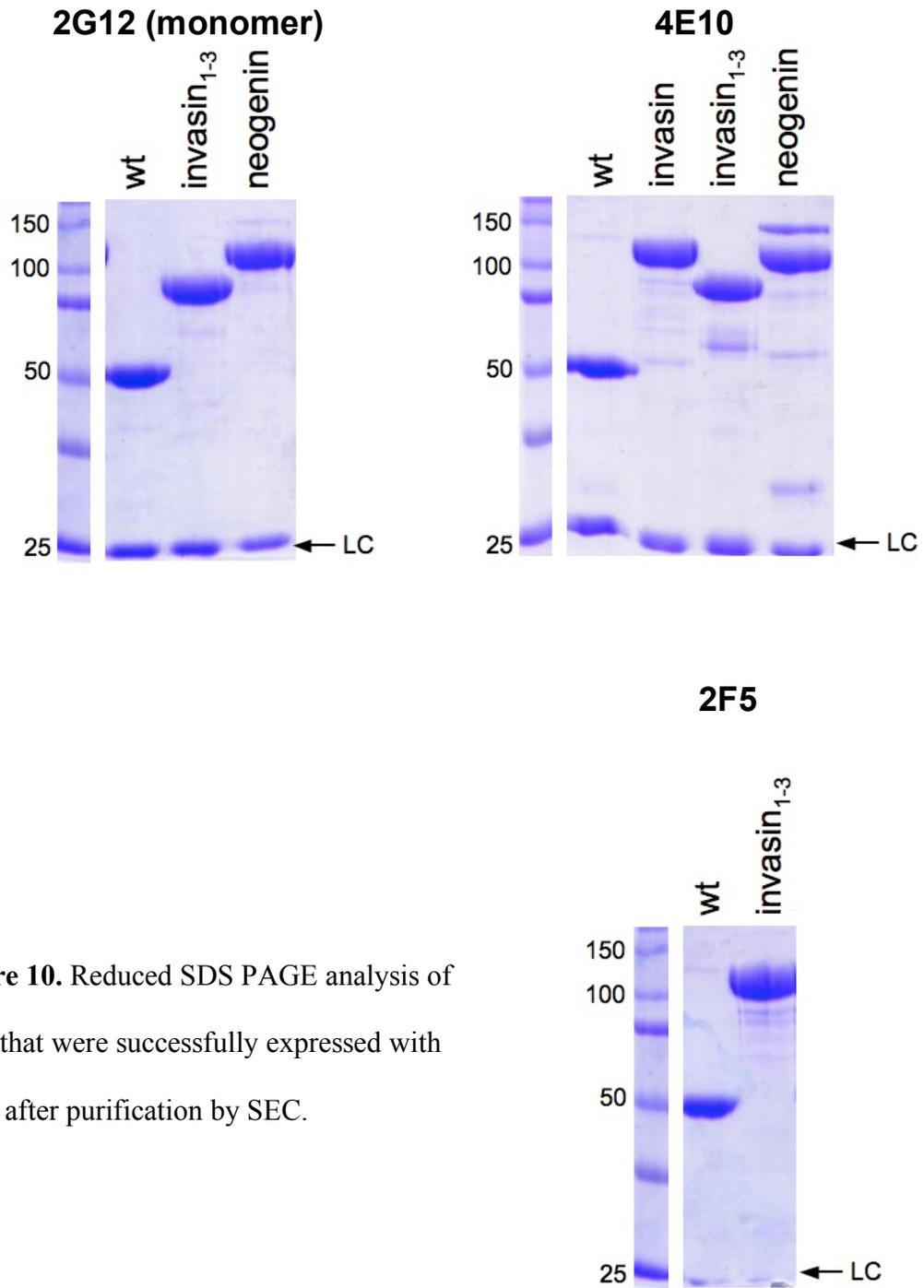
Figure 10. Reduced SDS PAGE analysis of IgGs that were successfully expressed with PFEs after purification by SEC.
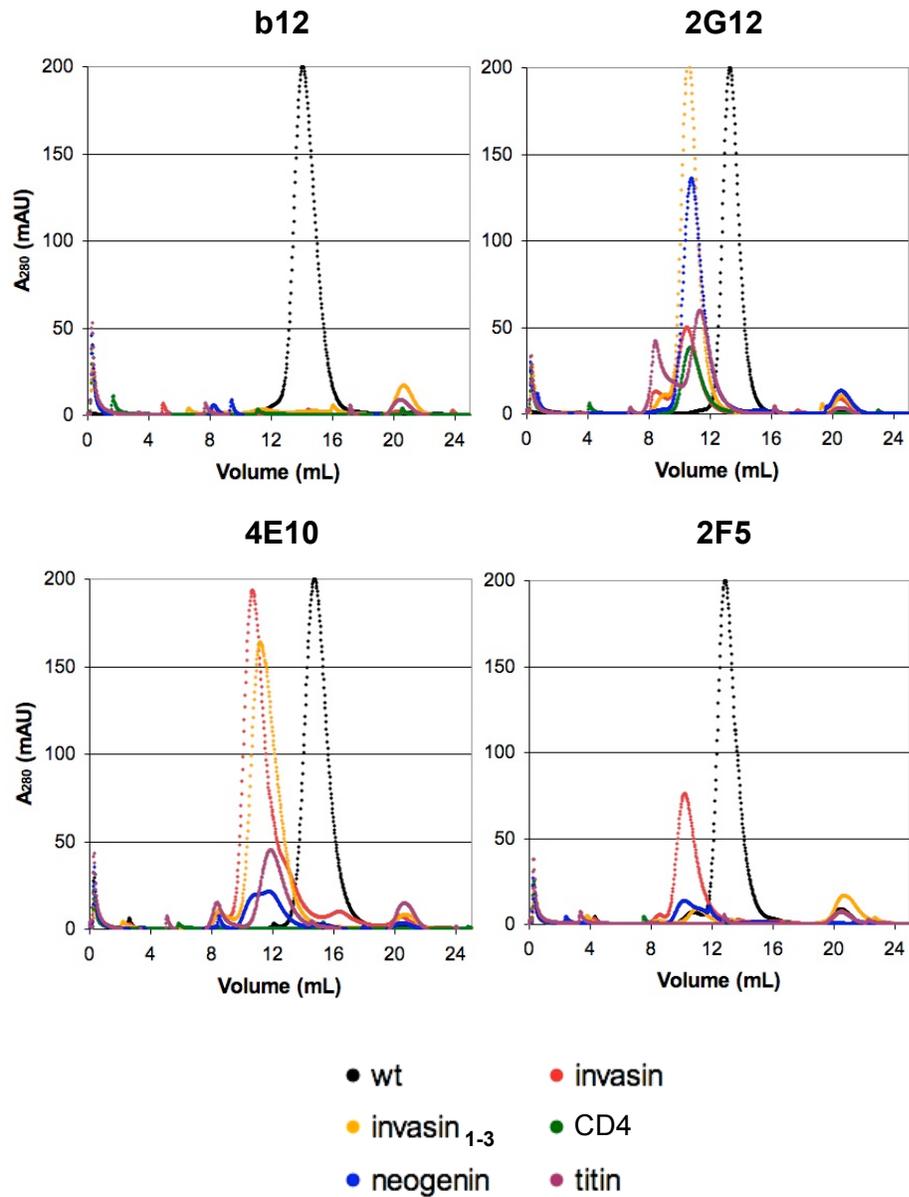
Figure 11. Gel filtrations profiles for IgGs with protein fusion extensions. Chromatograms for wt IgGs were normalized to peak heights of 200 mAU (~1 mg). The chromatograms for the hinge extended IgGs are shown untransformed such that differences in peak heights reflect differences in yield from 1 L of supernatant. The void peak is at 8.5 mL.
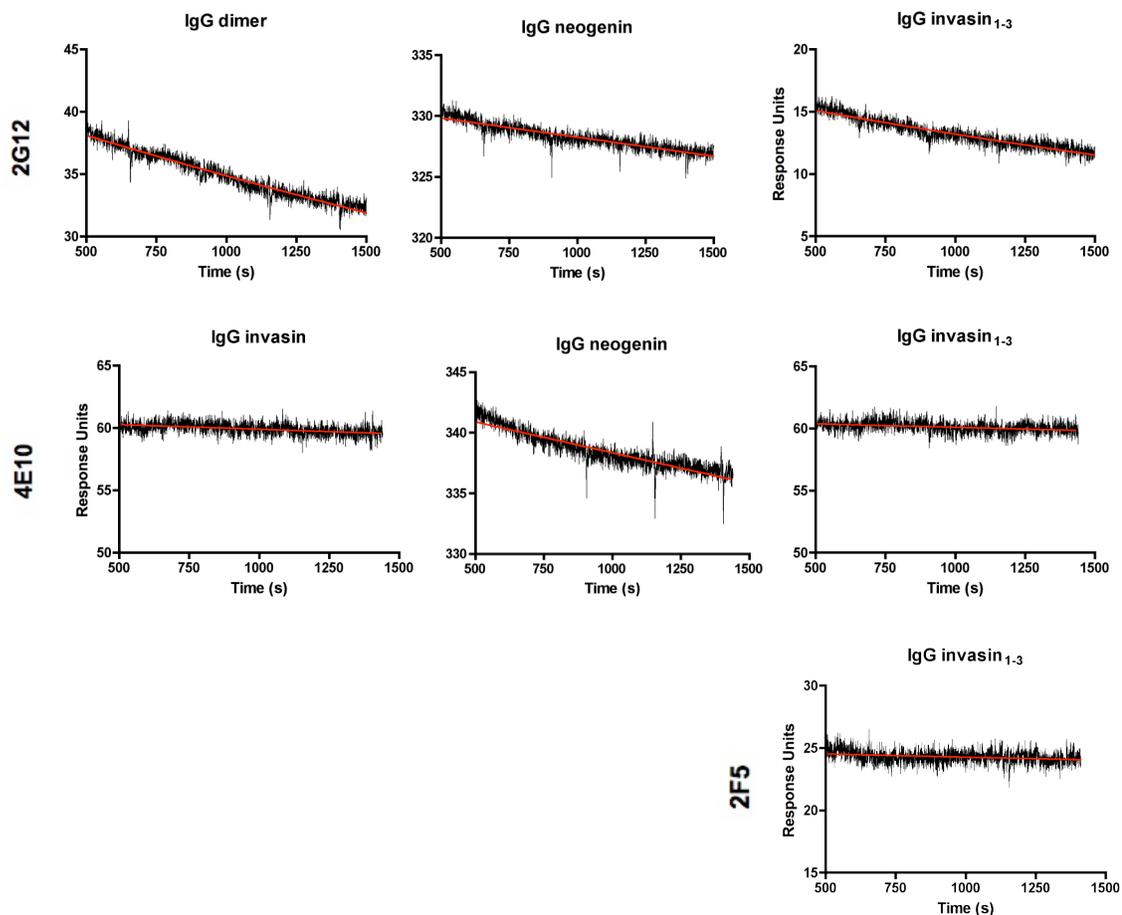
Figure 12. Dissociation curves for PFE constructs and 2G12 dimer bound to immobilized monomeric gp120 (2G12) or immobilized gp41 (4E10 and 2F5).

Table 4. Half-life times ($t_{1/2}$) calculated from the dissociation curves of dimeric IgG 2G12 and PFEs in Fig. 11.

Antibody	$t_{1/2}$ (min)	$t_{1/2}$ wt IgG : $t_{1/2}$ PFE
2G12 dimer	65	2.5
2G12 neogenin	1,200	0.14
2G12 invasin$_{1-3}$	43	3.8
4E10 invasin	920	1.1
4E10 neogenin	770	1.3
4E10 invasin$_{1-3}$	1,200	0.83
2F5 invasin$_{1-3}$	500	1.6
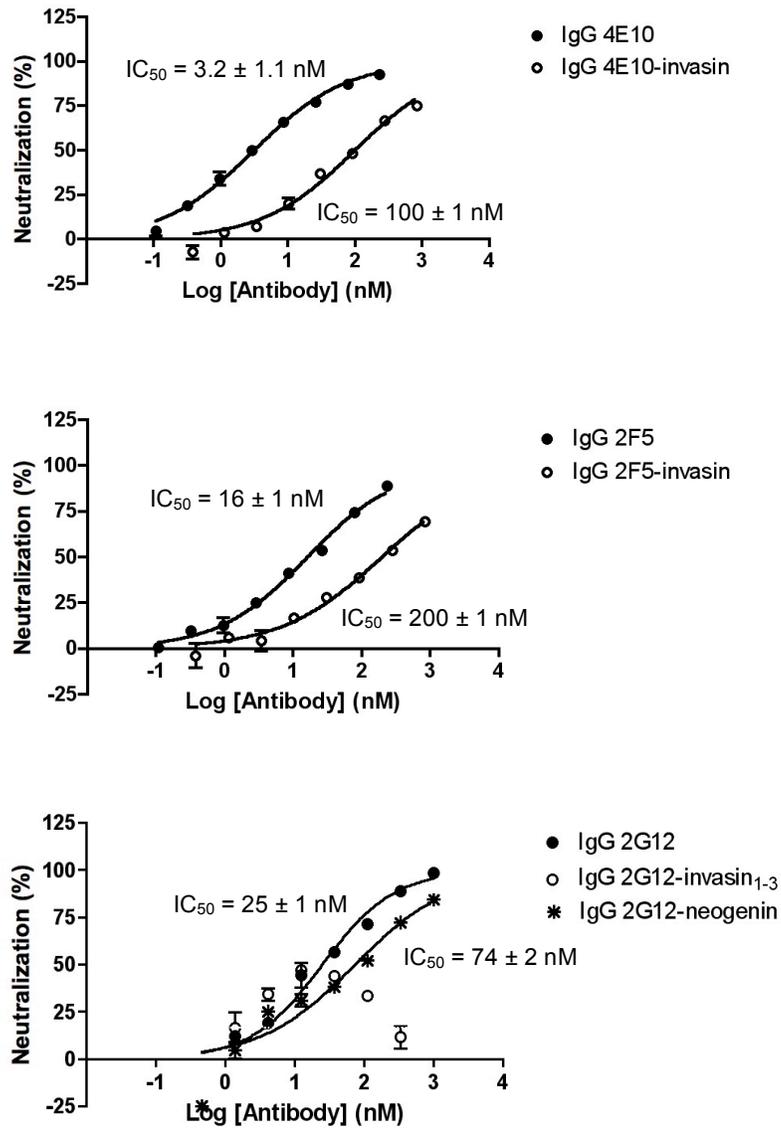
Figure 13. Comparison of neutralization activities for PFEs versus wild type IgGs (HIV strain 6535.3).

With respect to the 2G12 PFEs, no evidence was observed for increased neutralization potency. Given the increased size of the 2G12 PFEs and assuming that the domain swapped architecture of the 2G12 binding site is intact for these constructs, these results support the likelihood that the increased potency of dimeric IgG 2G12 derives from the presence of two adjoining domain swapped Fab dimers rather than the alternative possibility that the presence of two Fc domains increases its ability to sterically block the virus from binding to target cells. Thus, the reason for the lack of neutralization potency for purified 2G12 Fab and the precise mechanism by which this anti-carbohydrate antibody blocks infection remains unresolved.

The behavior of 2G12 with a tandem neogenin PFE was unexpected in that it could neutralize the virus at low concentrations with similar if not better potency as compared to monomeric IgG 2G12 (Fig. 13). However, above concentrations of ~20 nM, neutralization rapidly decreased to zero. Repeating this experiment yielded exactly the same result (data not shown). Additionally, injections of this construct over immobilized monomeric gp120 indicated an unexpectedly high maximum response (R_{max}) relative to monomeric IgG 2G12 that cannot be accounted for by its increased molecular weight ($R_{max} = C_{ligand} * S * MW_{analyte} / MW_{ligand}$ where R_{max} is the maximum response, C_{ligand} is the amount of immobilized ligand, and S is the stoichiometry of the interaction) (Fig. 12). Moreover, the calculated half-life for binding was ~7-fold slower than monomeric IgG 2G12 (1,200 min versus 160 min, Tables 3 and 4). A possible explanation for this behavior would be the presence of some affinity between neogenin and the 2G12 combining site such that as the neogenin PFE bound to gp120 on the chip surface, more binding sites were introduced, thereby permitting increased binding. Likewise, high concentrations in the neutralization

assay would enable the construct to act as a competitive inhibitor, rationalizing the drop in neutralization potency with increasing concentration of the inhibitor.

Discussion

The neutralization data for the RCEs indicate that extending the hinge region of an IgG using unstructured Gly$_4$Ser repeats does not enhance neutralization potency despite the observation in the binding analysis that the potential for avidity remains intact for all hinge extension lengths tested. There are at least three possible explanations for this observation:

(i) despite the low density of spikes on the surface of HIV, spikes can diffuse throughout the lipid bilayer at a rate that is not limiting to bivalent binding regardless of the distance between antibody combining sites;

(ii) the spikes are diffusion limited but the assumption that an unstructured linker enables the Fabs to efficiently sample extended conformations is false;

(iii) the spikes are diffusion limited and all spikes must be bound for neutralization to occur, but the set of nearest neighbor distances that describes the distribution of spikes for most virions leaves a significant portion of the spikes outside the reach of the antibodies with RCEs examined here.

The first explanation is not supported by the current results in which it was shown that a scBvFv architecture that has a maximum distance between the two combining sites of 11 nm is less potent than an IgG architecture where the maximum distance between

combining sites is approximately 15 nm (Chapter 2). For example, in the case of 4E10 it was observed that a scBvFv was equally potent to a monomeric single chain Fv, indicating a complete absence of cross-linking for the scBvFv even though evidence was observed of cross-linking for the IgG. Similarly, increased neutralization potency was observed for IgG b12 over the scBvFv architecture. Together, these data suggest that the distance between combining sites does influence neutralization potency and, consequently, the spikes on the surface of HIV are diffusion limited relative to the kinetics of bivalent binding.

An answer to the second explanation may be derived from a recent paper that compared the fluorescence resonance energy transfer (FRET) efficiency between GFP and YFP as a function of the number of repeats of a $(Gly_2Ser)_2$ linker connecting them (46). Here it was shown that the decrease in FRET efficiency as the linker length increased was consistent with a distribution function that shows a marked broadening in the range of distances sampled and that as each additional repeat of 6 amino acids was inserted, the average distance sampled increased by a decreasing fraction of the ~2 nm that 6 additional residues could theoretically span if fully extended. Consequently, rather than adding 15 nm, the $(Gly_4Ser)_5$ linker of the bivalent IgG (similar in length to CLY9 construct examined in the FRET investigation) probably added on average 4 to 5 nm to the 15 nm reach of the IgGs. An inability of the unstructured Gly_4Ser linkers used in this study to allow the Fabs to efficiently sample extended conformations could explain the lack of increased potency as a function of the number of linker repeats inserted into the hinge region.

The third explanation gives consideration to the stoichiometric requirements for neutralization and the nearest neighbor distribution reported for a set of 40 virions analyzed by electron microscopy (17). In a recent paper, researchers observed that influenza requires 8 to 9 functional spikes to mediate viral entry into a target cell but HIV and two other retroviruses tested required only a single functional spike to mediate successful entry (47). If this model is correct, then all or nearly all spikes must be bound for a single virus particle to be effectively neutralized if one assumes a high number of interactions between the virus particle and a target cell over the course of its lifetime. A counter to this argument was presented in another report in which HIV particles that appeared to be bound to target cells all exhibited what the researchers termed an "entry claw", in which ~5 spikes were involved in coordinating attachment (18). While this does not disprove the one-spike fusion hypothesis, it suggests that multiple spikes with intact CD4 and/or co-receptor binding sites (but not necessarily fusion-competent) are required to create a stable attachment prior to the fusion that may be catalyzed by a single spike.

Given the indications that (*i*) all or nearly all spikes must be bound either to prevent fusion or prevent a stable attachment, (*ii*) spikes are diffusion limited, (*iii*) RCEs are ineffective for expanding the reach of an IgG in this setting, and (*iv*) many HIV spike pairs are separated by distances that make them unrealistic targets for cross-linking, the design of an IgG capable of intra-spike cross-linking probably represents a more tractable goal. Because the distance between identical binding sites on a single spike is probably static, the effectiveness of this architecture would not be limited by potentially confounding factors such as spike mobility and nearest neighbor distribution.

If one considers how evolution appears to have selected for molecules with extended conformations, the linear arrangement of immunoglobulin (Ig) domains (*e.g.,* CD4, VEGF, PDGF) and fibronectin type III (FnIII) domains (*e.g.,* cell adhesion molecules) appear to be common themes compared to the evolution of unstructured linkers (*e.g.,* CD8). For example, a recent crystal structure of domains 168-170 from titin composed of two Ig domains and one FnIII domain revealed an extended conformation of ~12 nm (42), sufficient for the length requirement of b12. Unfortunately, measurable quantities of protein were not isolated for PFEs involving b12, the structure that was used to model and design the hinge extensions. Clearly, alternative approaches to developing structured linkers are required. Other approaches that may prove worthwhile include covalent modifications to couple Fabs *in vitro* using DNA molecules of different lengths or less complex PFEs such as small single domain proteins that are known for high yields and excellent stability.

The successful design of antibodies capable of intra-spike cross-linking could represent a significant advance in human health by providing the first realistic option for prophylaxis by gene therapy for HIV, or as in the case of chemical modification, applicability in a therapeutic setting. Moreover, perfection of a hinge extension technology developed to counteract the island effect that may be thwarting HIV neutralization by conventional antibodies might also be applied to other diseases or pathogens in which binding is avidity-limited.

Materials and Methods

Sequences for RCEs and PFEs were codon optimized and generated by DNA

synthesis. In the case of PFEs, a $(Gly_2Ser)_2$ linker was included at the N-terminus and C-terminus. All PFEs (invasin, $invasin_{1-3}$, CD4, $titin_{168-170}$, and a tandem repeat of $neogenin_{5-6}$) were further analyzed for predicted N-linked glycosylation sites that were subsequently removed with conservative Asn→Asp mutations or in a few cases, Ser/Thr→Ala mutations.

Full-length heavy chain (IgG1 subclass) and light chain (κ) genes were initially subcloned separately into pTT5. The RCEs and PFEs were then cloned immediately 5' to the natural IgG1 hinge sequence and 3' to the codon encoding the Cys residue that participates in the disulfide bond with light chain using the restriction sites NgoM IV and Nhe I previously introduced by PCR.

Methods describing expression, purification, binding assays, molecular modeling, and neutralization assays can be found in Chapter 2.

References

1.   2007 AIDS epidemic update : December 2007 (UNAIDS/WHO, November 2007).

2.   Hutter G*, et al.* (2009) Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* 360:692-698.

3.   Bedimo R*, et al.* (2006) Sustained HIV viral suppression following treatment interruption: an observational study. *AIDS Res Hum Retroviruses* 22:40-44.

4.   Roberts JD, Bebenek K, & Kunkel TA (1988) The accuracy of reverse transcriptase from HIV-1. *Science* 242:1171-1173.

5.   Trkola A*, et al.* (2005) Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat Med*. 11(6):615-622.

6.   Zwick MB*, et al.* (2001) Neutralization synergy of human immunodeficiency virus type 1 primary isolates by cocktails of broadly neutralizing antibodies. *J Virol* 75:12198-12208.

7.   Parren PW*, et al.* (2001) Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. *J Virol* 75:8340-8347.

8.   Lobo ED, Hansen RJ, & Balthasar JP (2004) Antibody pharmacokinetics and pharmacodynamics. *J Pharm Sci* 93:2645-2668.

9.   Pero JK, Haas EM, & Thompson NL (2006) Size dependence of protein diffusion very close to membrane surfaces: measurement by total internal reflection with fluorescence correlation spectroscopy. *J Phys Chem B* 110:10910-10918.

10.   Drake AW, Myszka DG, & Klakamp SL (2004) Characterizing high-affinity antigen/antibody complexes by kinetic- and equilibrium-based methods. *Anal Biochem* 328:35-43.

11.   Darbha R, *et al.* (2004) Crystal structure of the broadly cross-reactive HIV-1-neutralizing Fab X5 and fine mapping of its epitope. *Biochemistry* 43:1410-1417.

12.   Edwards MJ & Dimmock NJ (2001) A haemagglutinin (HA1)-specific FAb neutralizes influenza A virus by inhibiting fusion activity. *J Gen Virol* 82:1387-1395.

13.   Brunel FM, *et al.* (2006) Structure-function analysis of the epitope for 4E10, a broadly neutralizing human immunodeficiency virus type 1 antibody. *J Virol* 80:1680-1687.

14.   Martin-Garcia J, Cocklin S, Chaiken IM, & Gonzalez-Scarano F (2005) Interaction with CD4 and antibodies to CD4-induced epitopes of the envelope gp120 from a microglial cell-adapted human immunodeficiency virus type 1 isolate. *J Virol* 79:6703-6713.

15.   Wu H, *et al.* (2005) Ultra-potent antibodies against respiratory syncytial virus: effects of binding kinetics and binding valence on viral neutralization. *J Mol Biol* 350:126-144.

16.   Icenogle J, *et al.* (1983) Neutralization of poliovirus by a monoclonal antibody: kinetics and stoichiometry. *Virology* 127:412-425.

17.   Zhu P, *et al.* (2006) Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441:847-852.

18. Sougrat R, *et al.* (2007) Electron tomography of the contact between T cells and SIV/HIV-1: implications for viral entry. *PLoS Pathog* 3(5):571-581.

19. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, & Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109-113.

20. Chertova E, *et al.* (2002) Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), Is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J Virol* 76:5315-5325.

21. Schofield DJ, Stephenson JR, & Dimmock NJ (1997) Variations in the neutralizing and haemagglutination-inhibiting activities of five influenza A virus-specific IgGs and their antibody fragments. *J Gen Virol* 78 (Pt 10):2431-2439.

22. Karlsson Hedestam GB, *et al.* (2008) The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat Rev Microbiol* 6:143-155.

23. Pantaleo G & Koup RA (2004) Correlates of immune protection in HIV-1 infection: what we know, what we don't know, what we should know. *Nat Med* 10:806-810.

24. Plotkin SA (2001) Immunologic correlates of protection induced by vaccination. *Pediatr Infect Dis J* 20:63-75.

25. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, & Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455(7209):109-113.

26. Egan MA, Carruth LM, Rowell JF, Yu X, & Siliciano RF (1996) Human immunodeficiency virus type 1 envelope protein endocytosis mediated by a highly conserved intrinsic internalization signal in the cytoplasmic domain of

gp41 is suppressed in the presence of the Pr55gag precursor protein. *J Virol* 70:6547-6556.

27.   Ye L*, et al.* (2004) Surface stability and immunogenicity of the human immunodeficiency virus envelope glycoprotein: role of the cytoplasmic domain. *J Virol* 78:13409-13419.

28.   Benjamin J, Ganser-Pornillos BK, Tivol WF, Sundquist WI, & Jensen GJ (2005) Three-dimensional structure of HIV-1 virus-like particles by electron cryotomography. *J Mol Biol* 346:577-588.

29.   Yu X, Yuan X, Matsuda Z, Lee TH, & Essex M (1992) The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J Virol* 66:4966-4971.

30.   Rosenblatt S. http://www.tau.ac.il/lifesci/departments/biotech/members/rozenblatt/figures.html.

31.   Conway JF*, et al.* (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* 386:91-94.

32.   Yamaguchi M, Danev R, Nishiyama K, Sugawara K, & Nagayama K (2008) Zernike phase contrast electron microscopy of ice-embedded influenza A virus. *J Struct Biol* 162:271-276.

33.   Wei X*, et al.* (2003) Antibody neutralization and escape by HIV-1. *Nature* 422:307-312.

34.   Poignard P, Saphire EO, Parren PW, & Burton DR (2001) gp120: Biologic aspects of structural features. *Annu Rev Immunol* 19:253-274.

35.   Scanlan CN, *et al.* (2002) The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. *J Virol* 76:7306-7321.

36.   Lin G, *et al.* (2003) Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. *J Virol* 77:1337-1346.

37.   Calarese DA, *et al.* (2003) Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 300:2065-2071.

38.   Binley JM, *et al.* (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J Virol* 78:13232-13252.

39.   Bird RE, *et al.* (1988) Single-chain antigen-binding proteins. *Science* 242:423-426.

40.   Wu H, Kwong PD, & Hendrickson WA (1997) Dimeric association and segmental variability in the structure of human CD4. *Nature* 387:527-530.

41.   Hamburger ZA, Brown MS, Isberg RR, & Bjorkman PJ (1999) Crystal structure of invasin: a bacterial integrin-binding protein. *Science* 286:291-295.

42.   Mrosek M, *et al.* (2007) Molecular determinants for the recruitment of the ubiquitin-ligase MuRF-1 onto M-line titin. *FASEB J* 21:1383-1392.

43.   Cai Y, *et al.* (2008) Optimizing the codon usage of synthetic gene with QPSO algorithm. *J Theor Biol* 254:123-127.

44.   Sandhu KS, Pandey S, Maiti S, & Pillai B (2008) GASCO: genetic algorithm simulation for codon optimization. *In Silico Biol* 8:187-192.

45.   Shewmake TA, Solis FJ, Gillies RJ, & Caplan MR (2008) Effects of linker length

       and flexibility on multivalent targeting. *Biomacromolecules* 9:3057-3064.

46.   Evers TH, van Dongen EM, Faesen AC, Meijer EW, & Merkx M (2006)

       Quantitative understanding of the energy transfer between fluorescent proteins

       connected via flexible peptide linkers. *Biochemistry* 45:13183-13192.

47.   Yang X, Kurteva S, Ren X, Lee S, & Sodroski J (2005) Stoichiometry of

       envelope glycoprotein trimers in the entry of human immunodeficiency virus type

       1. *J Virol* 79:12132-12147.

CHAPTER 4

Potent neutralization of HIV is not a predictor of the ability

to trigger antibody-dependent cellular cytotoxicity

This work was completed in collaboration with my colleague, Alexandre Webster, with assistance from Priyanthi Gnanapragasm and Jost Vielmetter. Using a panel of anti-HIV antibodies, it was observed that their respective abilities to trigger antibody dependent cellular cytotoxicity varied from highly effective to completely ineffective despite the fact that each of the antibodies exhibited potent neutralization of a virus of the same strain as used for the cytotoxicity assay.

Introduction

Certain classes of immunoglobulins can prevent or contribute to the elimination of a viral infection by other mechanisms in addition to neutralization (*i.e.,* binding directly to virus particles to prevent entry into target cells). These additional mechanisms include antibody-dependent cellular cytotoxicity (ADCC), phagocytosis, and activation of the "classical" pathway of complement. With respect to the humoral immune response, the cell-mediated responses of ADCC and/or phagocytosis can be triggered by the interaction of antibody bound to antigen on an infected cell and Fc receptors expressed on the surfaces of various types of innate immune effector cells including natural killer (NK) cells, macrophages, neutrophils, and subsets of γδ T cells (1, 2). Complement activation is triggered by the binding of soluble factors in the plasma to antibody-antigen complexes, beginning with the fixation of C1q, which activates a cascade of reactions that culminate in the formation of pores resulting in lysis of infected cells as well as direct viral inactivation (1, 3). Each of these mechanisms has been shown to play significant roles in controlling and/or eliminating various pathogens (4), but relative to the CTL response and neutralization, the contributions of these responses to the control of HIV replication or their potential value to the prevention of infection remains largely uncharacterized.

Mechanism of ADCC

Of the four subclasses of IgG, only subclasses 1 and 3 are able to trigger ADCC activity in humans (1), which occurs via cross-linking CD16 (FcγRIII), a low micromolar affinity activating Fc receptor (5). CD16, which is a 50-80 kDa highly glycosylated

protein, is expressed as two isoforms: a transmembrane anchored form with a cytoplasmic domain (CD16a) and a glycophosphatidylinositol (GPI) linked form (CD16b) (6). Whereas the precise role of CD16b is not yet fully understood, the function of CD16a on NK cells and macrophages in mediating ADCC is well established (2, 6, 7). Cross-linking of CD16a by antibodies serves to increase the concentrations of the intracellular immunoreceptor tyrosine-based activations motifs (ITAMs) of heterodimers and/or homodimers that belong to the CD16a-associated signaling chains, FcεRIγ and CD3ζ (8), and the resulting protein tyrosine kinase cascade leads to granule exocytosis.

Apoptosis, or programmed cell death, is induced during ADCC following a similar series of steps that describe the mechanism of CTL attack: cytotoxic granules are released from the cytosol of effector cells containing perforin, granzymes, and granulysin (1). Perforin acts in a similar manner to the C9 component of the complement cascade in that it creates pores in the cytoplasmic membrane of the target cell, and granzymes, which then gain entry to the cytosol of the target cell, activate the caspase cascade, culminating in apoptosis (1). However, whereas individual CTLs are specific for individual peptide-MHC class I complexes presented on the surfaces of cells, effector cells that mediate ADCC recognize the Fc domains of bound IgG1 or IgG3 and are therefore competent to attack any infected cell that is expressing a recognized antigen on its surface.

Evidence for a contribution of ADCC to HIV control

Given the lack of evidence that neutralizing antibodies significantly contribute to the control of viral replication during the natural course of infection, researchers are

turning their attention to the role that the humoral immune response might play in recruiting Fc receptor-dependent mechanisms of eliminating infected cells (3, 9-12). This interest has been at least partly fueled by the observations that some antibodies that cannot neutralize virus can still recruit these mechanisms and that these antibodies may appear well before the development of neutralizing antibodies (3). In addition, it has been reported that the ADCC activity titer of sera collected from SIV challenged NHPs vaccinated against SIV is inversely correlated with viral titer and directly correlated with $CD4^+$ T-cell concentrations (10).

Further underscoring the need to carefully examine the contributions of these mechanisms to the control of HIV, ADCC was recently shown to be critical to the ability of IgG1 b12 to protect NHPs from viral challenge (12). When this study was first carried out with the observation that a dosage of 25 mg/kg was necessary to completely protect animals challenged with $SHIV_{SF162P3}$ (13), the conclusion was understood at the time to be that b12 had fully neutralized the virus at this dosage. However, when researchers repeated the study six years later but instead made two different mutants of b12, one that could not fix complement and another than could not bind CD16, the dosage of 25 mg/kg was no longer sufficient for the non-CD16-binding mutant to protect the animals (12). Thus, even when more than 5% of the total IgG in the blood is composed of a monoclonal antibody to which the challenge virus is unusually sensitive, it can still not be enough to fully protect from viral challenge. Further elucidation of the potential contribution of antibodies to viral control by mechanisms other than neutralization may eventually alter the currently accepted view that the humoral immune response is negligible relative to the CTL response in controlling the natural course of infection.

The results presented here are directed towards asking (*i*) if the ability to potently neutralize HIV necessarily implies the ability to efficiently recruit ADCC and (*ii*) if antibodies directed against one region of the envelope spike versus another might result in differing abilities to trigger this mechanism. If ADCC is a critical component to protection, then the possibility that various epitope classes might differ with respect to the ability of antibodies that are directed against them to recruit ADCC should be examined. For example, anti-CD4 binding site (anti-CD4bs) antibodies might be better able to recruit ADCC than antibodies that bind the MPER as a result of greater accessibility to the Fc domain by CD16a. Towards this end, an *in vitro* ADCC assay was developed in which ^{35}S-labeled target cells stably expressing envelope spikes are incubated with antibody and peripheral blood mononuclear cells (PBMCs) and then examined for the release of ^{35}S into the media. The abilities of a panel of bNAbs to neutralize HIV and trigger ADCC were compared, and the implications of these results on vaccine design are discussed.

Results

Selection of a gp160-positive cell line. Three different cell lines that were putatively positive for stable expression of gp160 were obtained from the NIH AIDS Research and Reference Reagent Program (CHO-gp160 (14), HeLa 69T1RevEnv (15), and Jurkat 7/3-STOP (16)), and a fourth was kindly provided by Dr. Pin Wang (USC). Flow cytometry analysis of each of the cell lines stained with FITC-conjugated polyclonal goat anti-gp120 antibody revealed only one cell line to be uniformly expressing gp160: CHO-gp160 (Fig. 1). This cell line was then selected as the basis for the development of the
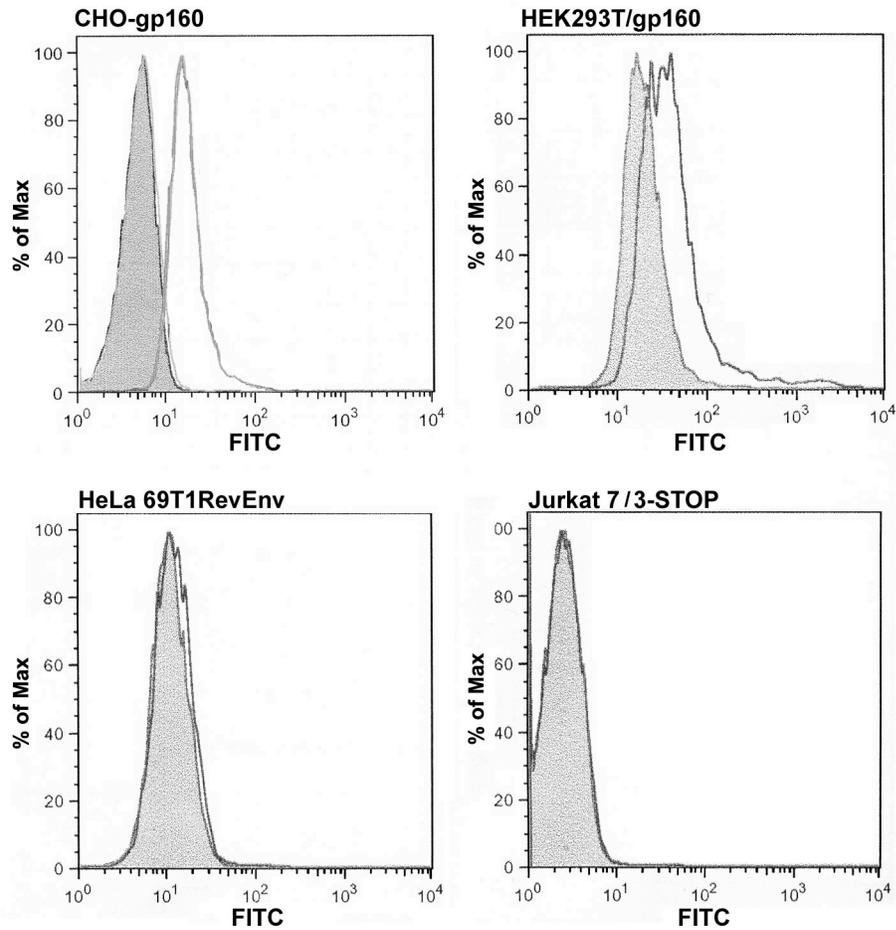
Figure 1. Four cell lines examined for stable surface expression of gp160. Cells were stained with 1 µg/mL FITC-conjugated polyclonal goat anti-gp120 antibody and examined by flow cytometry analysis (grey, unstained; clear, stained).

ADCC assay. A panel of broadly neutralizing IgG1 antibodies including 2G12 monomer, 2G12 dimer, b12, 447-52D, 4E10, and 2F5 were then evaluated for their abilities to neutralize strain HxBc2 (also called IIIB), a T-cell line adapted clade B strain of HIV that is particularly sensitive to each of these antibodies. The antibodies were then evaluated for their abilities to trigger ADCC against CHO-gp160.

Conditions for the ADCC assay included a 30-minute pre-incubation period of antibody with target cells on ice followed by the addition of human PBMCs and incubation for 4 hours at 37 °C at an effector-to-target cell ratio of 50:1. The addition of a monoclonal anti-CD20 isotype control antibody, Rituxan, did not result in non-specific lysis (Fig. 2). In all cases where ADCC activity was observed, the percentage of specific lysis either began to level off or decrease at concentrations above ~10 µg/mL (Fig. 2), probably reflecting increased competition for CD16a between bound and unbound IgG1.

Evaluation of anti-gp120 monoclonal antibodies. Monomeric 2G12 was observed to be effective at recruiting ADCC *in vitro* with a half maximal effective dose (ED_{50}) of 1.1 µg/mL, which was comparable to its 50% inhibitory concentration (IC_{50}) for the *in vitro* neutralization assay of 0.36 µg/mL (Table 1). In agreement with previous results (Appendix F), dimeric IgG1 2G12 was observed to be more potent than the monomer, with ED_{50} and IC_{50} values of 0.32 µg/mL and 0.032 µg/mL, respectively. The performance of these anti-carbohydrate antibodies are unlikely to have arisen from non-specific reactivity as no significant specific lysis was observed when tested against the untransfected parental cell line, CHO K1 (Fig. 3).

The anti-CD4bs antibody b12 exhibited potent ED_{50} and IC_{50} values of 0.058 μg/mL and 0.033 μg/mL, respectively (although the low Hill coefficient of the curve precluded an accurate determination of the ED_{50} value) (Table 1 and Fig. 2). However, the maximal lysis observed for b12 was only 23%, significantly lower than that observed for monomeric 2G12 or dimeric 2G12 (75% and 63%, respectively). It is unlikely that this weak maximal lysis was due to antibody-induced shedding of gp120 as almost no detectable shedding was observed in the presence of b12 and other CD4bs antibodies from the surface of HIV-infected cells in a previous study whereas soluble CD4 (sCD4) was shown to induce the release of over 5 μg of gp120 in 2 hours under identical conditions (17).

The anti-V3-loop antibody, 447-52D, exhibited ED_{50} and IC_{50} values of 0.20 μg/mL and 0.083 μg/mL, respectively. As compared to b12, a similar result was observed in the ADCC assay for 447-52D, with a low maximal lysis value (22%). However, unlike the case for anti-CD4bs antibodies, anti-V3 loop antibodies have been shown to efficiently induce gp120 shedding at rates that are nearly comparable to sCD4 for T-cell line adapted strains of HIV (17), rationalizing its poor performance in the ADCC assay.

The anti-MPER antibodies 4E10 and 2F5 do not exhibit ADCC activity. By contrast to the anti-gp120 antibodies, 4E10 and 2F5 were unable to elicit detectable ADCC activity at any of the concentrations tested even though both were extremely potent in virus neutralization (Fig. 2), yielding IC_{50} values of 0.064 μg/mL and 0.026 μg/mL, respectively.
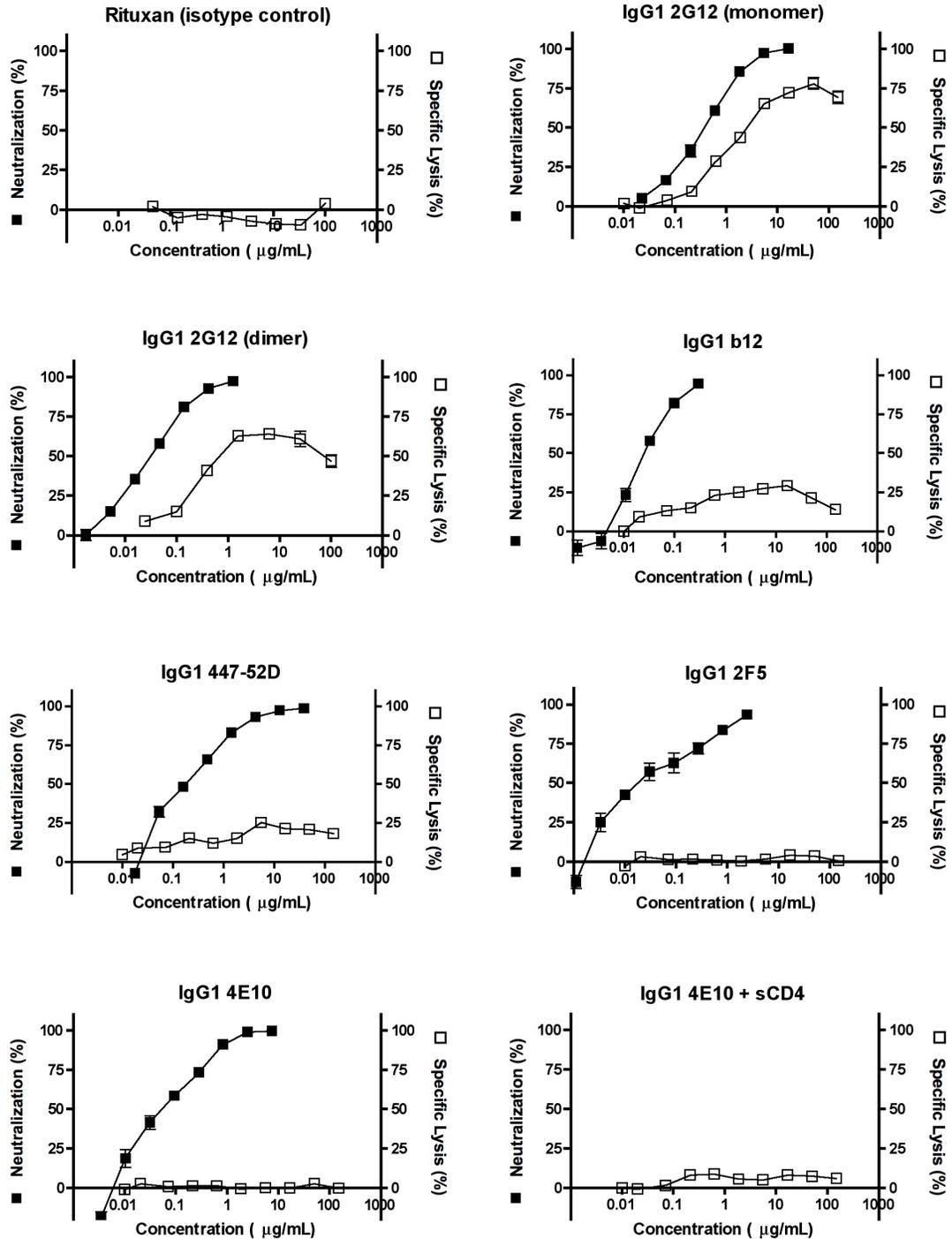
Figure 2. Summary of curves for the *in vitro* neutralization data (■) and *in vitro* ADCC data (□).
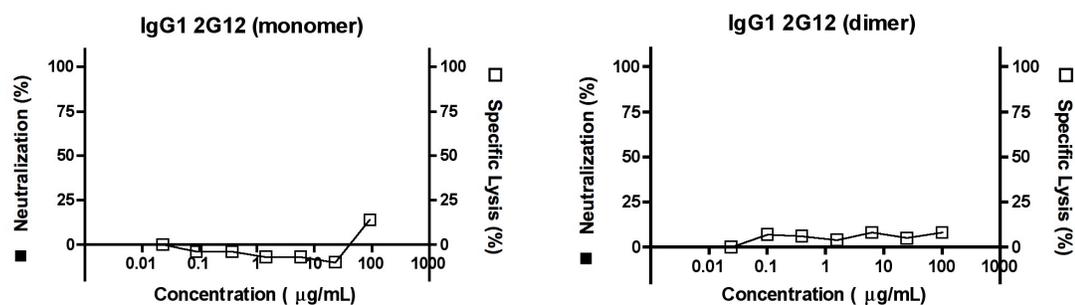
Figure 3. Summary of curves from the *in vitro* ADCC data for control tests of IgG1 2G12 monomer and dimer versus untransfected CHO cells (□).

Table 1. Summary of calculated 50% effective doses (ED$_{50}$) for specific lysis in the *in vitro* ADCC assay and 50% inhibitory concentrations (IC$_{50}$) for the *in vitro* neutralization assay (n.d., not done).

Antibody	ADCC EC$_{50}$ (µg/mL)		Max lysis	Neutralization IC$_{50}$ (µg/mL)		
Rituxan (IgG1 isotype control)	No activity			n.d.		
IgG1 2G12 (monomer)	1.1	± 0.2	75%	0.36	±	0.11
IgG1 2G12 (dimer)	0.32	± 0.05	63%	0.032	±	0.007
IgG1 b12	0.058	± 0.125 *	23%	0.033	±	0.014
IgG1 447-52D	0.083	± 2.6 *	22%	0.20	±	0.04
IgG1 4E10	No activity			0.064	±	0.017
IgG1 4E10 + sCD4	No activity			n.d.		
IgG1 2F5	No activity			0.026	±	0.010

* The weak ADCC activity demonstrated in these curves prevented an accurate determination of their ED$_{50}$ values.

Some data suggest that the addition of sCD4 may potentiate access to the 4E10 epitope (18), but when sCD4 was added at an equimolar concentration to 4E10, no significant increase in activity was observed (Fig. 2), indicating that anti-MPER antibodies in general may be unable to trigger ADCC.

Discussion

In this investigation, the performance of a panel of bNAbs was compared using an *in vitro* neutralization assay and a newly developed *in vitro* ADCC assay. All of the antibodies were potent neutralizers of the strain of virus used (HXBc2), with each one able to neutralize ~100% of the virus at concentrations at or below ~10 μg/mL. However, the antibodies exhibited wide variation in their abilities to recruit ADCC activity.

Monomeric and dimeric IgG1 2G12 were particularly effective at recruiting ADCC, yielding typical sigmoidal dose-response curves and maximal specific lysis values of 75% and 63%, respectively. Importantly, these data confirmed that the Fc domains of dimeric IgG1 2G12 can bind CD16 and efficiently recruit ADCC with enhanced potency over monomeric IgG1 2G12. By contrast, IgG1 b12 and IgG1 447-52D exhibited only modest abilities to recruit ADCC. However, perhaps most striking was the absence of any detectable ADCC activity upon the addition of IgG1 4E10 or 2F5, which target adjacent epitopes of the MPER at the base of the stalk of the trimeric envelope spike. Even with the addition of sCD4, which has been shown to potentiate neutralization by 4E10 (19), presumably by promoting access to its epitope or expanding the time period during which it is available for binding, no detectable ADCC activity was observed.

One explanation that may help to rationalize the observed differences in ADCC activity is the requirement for access by CD16a, a single pass transmembrane protein composed of two immunoglobulin domains that adopt a U-shaped conformation (20, 21). An atomic resolution X-ray crystal structure of the extracellular portion of CD16a in complex with human Fc shows that the CD16a binding site is located on the upper portion of the CH_2 domain (20), which likely places constraints on the rotational freedom of the Fabs (Fig. 4). But perhaps more importantly, the C-terminal sequence that links the second domain of CD16a to the transmembrane sequence is 16 residues long (2). Thus, in the unlikely event of adopting a fully extended conformation, the base of the second domain of CD16a is limited to a separation distance from the effector cell's membrane of ~6.5 nm (Fig. 5).

Given that the trimeric envelope spike extends ~12 nm from the membrane and the lobes created by the gp120 moieties extend several nanometers from the stalk (22), binding to different epitopes may impose demands on CD16a accessibility that cannot be met in some cases (Fig. 5). For example, whereas 2G12 may be binding one or more glycans on the upper portion of spike (23), and b12 binds roughly parallel to the spike (24), anti-MPER antibodies such as 4E10 and 2F5 bind near the base (25, 26). Models of these binding modes indicate that even a fully extended CD16a receptor might have difficulty extending far enough to reach an Fc domain when the antibody is targeting the stalk portion of the envelope spike (Fig. 5), particularly in the presence membrane proteins that extend out from the surface of a typical cell (Fig. 5*D*). Consistent with this hypothesis, in the original paper where 4E10 and 2F5 were first characterized (along with many other antibodies), only 4E10 and 2F5 were reported to derive from IgG3 subtypes (27).
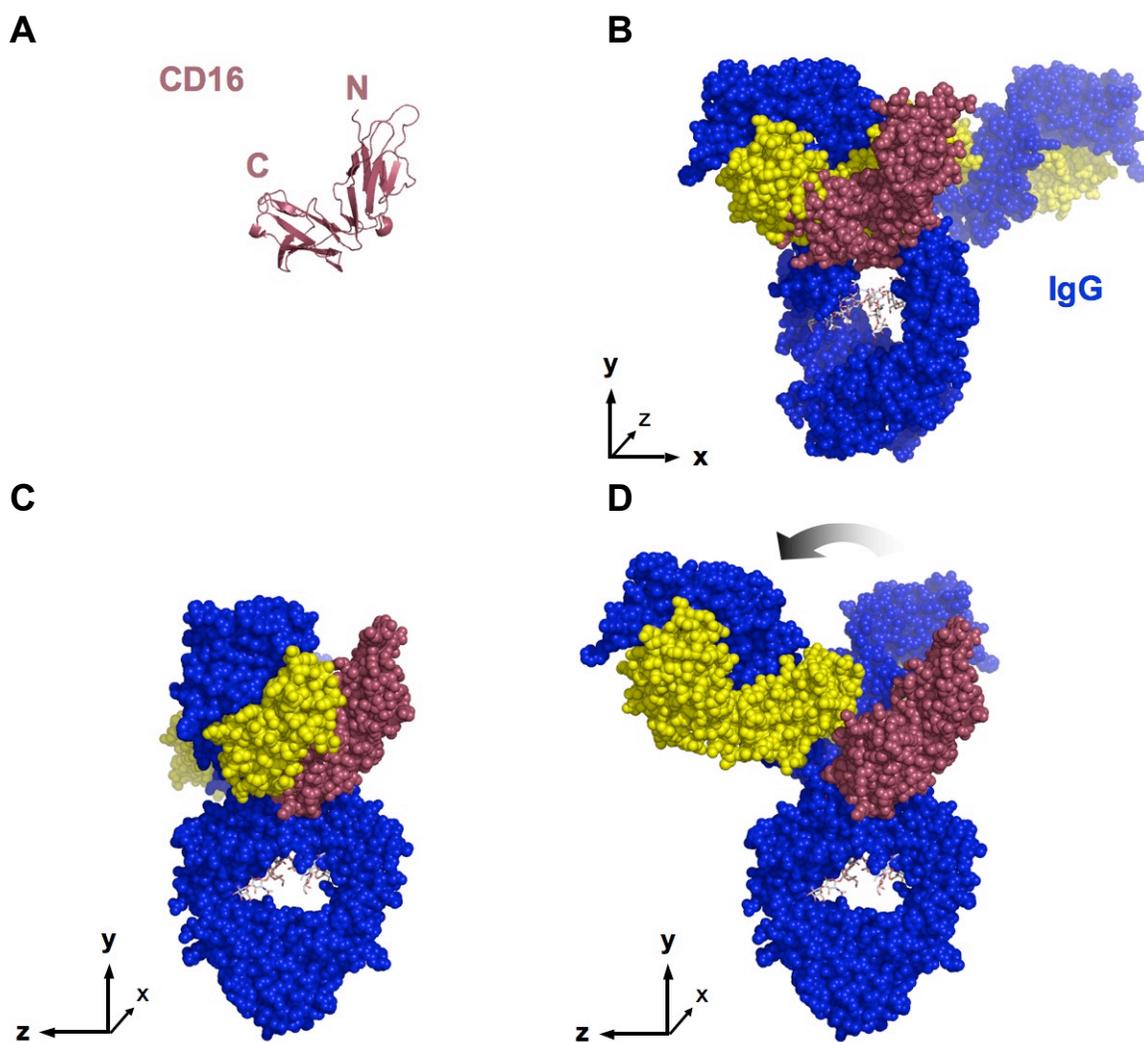
Figure 4. Model structure of IgG-CD16 complex. (*A*) Ribbon diagram for the X-ray crystal structure of CD16 extracellular domain in complex with human Fc (not shown) (PDB ID code 1E4K (20)). (*B*) X-ray crystal structure of IgG1 b12 (PDB ID code 1HZH (28)) docked to CD16. The Fc domains from the two PDB files were used for alignment. (*C*) Same as figure *B* but rotated 90°, showing simultaneous occupancy of space between the Fabs and CD16 when the structures are unaltered. (*D*) Same as figure *C* but the Fabs have been rotated to eliminate clashes between CD16 and the antibody, demonstrating that CD16 binding may impose constraints on Fab flexibility.
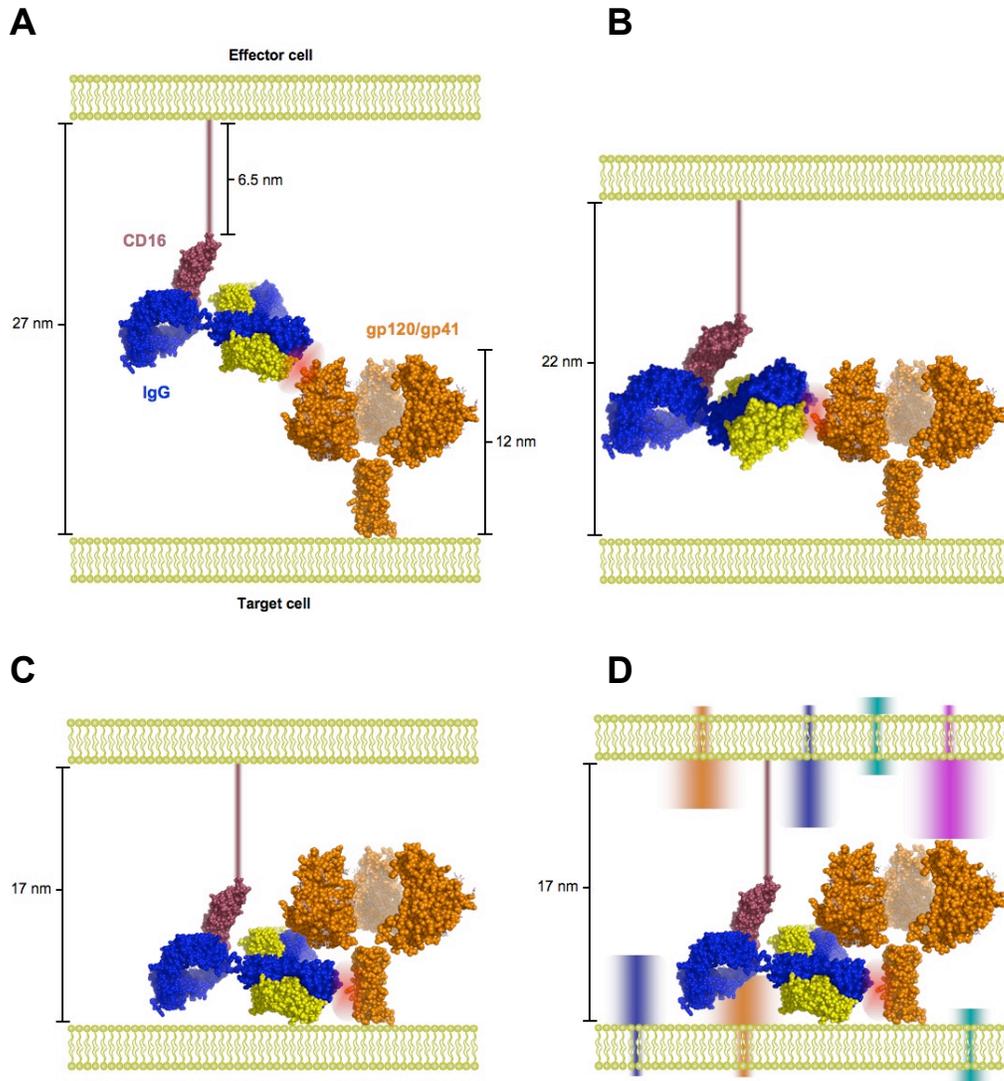
Figure 5. Conceptual model of the distance requirements for simultaneous binding of IgG1 with CD16 and the HIV envelope spike with the assumption that the 16-residue unstructured linker connecting the extracellular domains of CD16 to the membrane is adopting a fully extended state. The height for the envelope spike was based on reports of tomographic reconstructions of intact spikes (22). Schematics for antibody binding to (*A*) glycan clusters or the V3 loop region near the top of the spike, (*B*) the CD4 binding site, and (*C*) the MPER. (*D*) An illustration of the potential steric effects imposed by the presence of other membrane-bound proteins.

 The architecture of IgG3 differs significantly from IgG1 in that it has an unusually long hinge region that serves to increase the distance between the Fc domain and the Fabs. Thus, it may be that the Fc domains of anti-gp120 antibodies are inherently more accessible to CD16a binding as compared to anti-MPER antibodies.

An additional explanation for the exceptional performance of the 2G12 antibodies in the ADCC assay relative to the antibodies that bind protein epitopes may be that multiple 2G12 antibodies are binding to each trimeric spike with avidity as a result of its high glycan content (see Chapter 3), thereby inducing localized high concentrations of bound CD16a on effector cells and amplifying the ADCC activation signal. For protein epitopes, there can exist no more than three binding sites per envelope spike, which cannot be cross-linked by a single antibody due to geometric constraints (see Fig. S3, Chapter 2). Consequently, the endocytosis signals in the cytoplasmic tail of gp41 that serve to limit the concentration of spikes on the surface of a cell (29, 30) may present a limit to highly avid binding by these antibodies and limit their capacity to cross-link CD16a.

The finding that CD16-binding was critical to the ability of IgG1 b12 to protect NHPs in the SHIV challenge model (12) highlights the possibility that future advancements in optimizing the therapeutic efficacy of anti-HIV antibodies may depend on their abilities to trigger ADCC. The clinical trials discussed in Chapter 3 where a cocktail composed of 2G12, 4E10, and 2F5 suppressed viral rebound among patients who underwent interruptions in HAART also demonstrated that escape mutations were found only for 2G12 and not 4E10 or 2F5 (31). Thus, it is tempting to speculate that the

selective pressure of 2G12 might have derived, at least in part, from an enhanced ability to trigger ADCC against infected cells.

Given the nearly universal failure to elicit potent neutralizing antibodies against HIV, one possible approach to increasing the effectiveness of the humoral immune response would be to develop a method to introduce engineered variants of anti-HIV antibodies with increased abilities to trigger ADCC by gene therapy. For example, Fc mutations have been identified that increase the affinity of the Fc for CD16a by greater than 2 orders of magnitude (32), resulting in a similar increase in ADCC activity *in vitro* as well as increased potency *in vivo*. As an additional approach, we are currently examining whether the RCEs described in Chapter 3 might exhibit increased ADCC activity as a result of increasing the hinge length.

Methods

Untransfected CHO-K1 cells were obtained from the American Type Culture Collection (ATCC) and maintained in R-10 media (RPMI-1640 supplemented with 10% fetal calf serum (FCS), 1% L-glutamine, 1% 5000 U/mL Pen/Strep, 1% NEAA, and 1% Na-pyruvate). The HxBc2 gp160$^+$ CHO cell line, derived from CHO-K1 cells and named CHO-WT (referred to as GHO-gp160 in this text) was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from C. Weiss and J. White (14). These cells were maintained in G-MEM with no glutamine and supplemented with 10% dialyzed FCS, 3.7% $NaHCO_3$ at 75 mg/mL, 1% Pen/Strep at 5000 U/mL, 1% 100x non-essential amino acids, 1% 100x Na-pyruvate, 1% 100x glutamate/asparagine (each at 6 mg/mL), 2% 50x Nucleosides (3.5 mg/mL of cytidine,

uridine, guanosine, and adenosine; 1.2 mg/mL of thymidine), and 0.4% methionine sulfoximine (MSX) at 18 mg/mL (selection for glutamine synthetase transfectants).

PBMCs were purified from a lymphapheresis pack supplied by a single donor (Hemacare) as described in the manual (Ficoll Paque Plus User Manual, GE Healthcare). Briefly, donor blood was combined 1:1 with RPMI-1640 and 30 mL aliquots were layered onto 15 mL aliquots of Ficoll-Paque Plus in 50 mL conical tubes and centrifuged at 300 x g for 30 minutes at 4 °C. Buffy coats were then washed three times with RPMI 1640 and re-suspended in FCS to a concentration of 200 million cells per mL. 100 million cells were then combined 1:1 with freezing medium (20% DMSO, 80% FCS), frozen at -80 ºC, and stored in liquid nitrogen. PBMCs were reconstituted in R-10 media 24 to 48 hours prior to use in the assay.

Each antibody reagent was tested in triplicate. One million target cells were seeded into a six well plate and allowed to adhere overnight. The growth media were then replaced with identical media but lacking Glu/Met/Cys and supplemented with S^{35}-Met to 0.25 mCi. After incubation overnight at 37ºC, the cells were then detached with 2 mM EDTA in PBS, washed three times with R-10, and resuspended to a concentration of 200,000 cells per mL in R-10. Ten thousand target cells were then dispensed per well, to which 50 µL of antibody diluted into R-10 were added and incubated for 30 minutes on ice. PBMCs were washed once in R-10 and added to the appropriate wells to a final volume of 200 µL per well (100 µL/well for the maximum lysis condition). Plates were incubated for 4 hours at 37 ºC. Fifteen minutes prior to the end of the incubation period, 100 µL of 1 M NaOH was added to wells for the maximum lysis condition containing only labeled target cells. Plates were then centrifuged at 300 x g for 5 minutes and 50 µL

of each sample was transferred to white-bottomed Optiplate-96 reader plates already containing 200 µL of Microscint 40 scintillation fluid (PerkinElmer). Plates were then wrapped in foil and allowed to sit overnight at room temperature. Samples were read in a Wallac 1450 Microbeta TriLux. Specific lysis was calculated as $[(\alpha - \beta)/(\gamma - \beta)]$ x 100%, where α is target cells with PBMCs and antibody, β is maximum lysis, and γ is target cells with PBMCs alone. The curves were fit with a four-parameter logistic equation (Prism 4 for Macintosh, GraphPad Software, Inc.).

The FITC-conjugated polyclonal goat anti-gp120 antibody was obtained from Abcam. Methods describing expression, purification, binding assays, neutralization assays, and molecular modeling can be found in Chapter 2. Rituxan was generously provided by Dr. Sanjeev Nandakumaran (Kaiser Permanente Southern California Medical Group).

References

1.   Janeway CA, Travers P, Walport M, & Schlomchik MJ (2005) *Immunobiology* (Garland Science Publishing, New York, NY) 6th Ed.

2.   Rogers KA, Scinicariello F, & Attanasio R (2006) IgG Fc receptor III homologues in nonhuman primate species: genetic characterization and ligand interactions. *J Immunol* 177:3848-3856.

3.   Aasa-Chapman MM*, et al.* (2005) Detection of antibody-dependent complement-mediated inactivation of both autologous and heterologous virus in primary human immunodeficiency virus type 1 infection. *J Virol* 79:2823-2830.

4.   Burton DR (2002) Antibodies, viruses and vaccines. *Nat Rev Immunol* 2:706-713.

5.   Simmons D & Seed B (1988) The Fc gamma receptor of natural killer cells is a phospholipid-linked membrane protein. *Nature* 333:568-570.

6.   Nagarajan S*, et al.* (1995) Ligand binding and phagocytosis by CD16 (Fc gamma receptor III) isoforms. Phagocytic signaling by associated zeta and gamma subunits in Chinese hamster ovary cells. *J Biol Chem* 270:25762-25770.

7.   Mandelboim O*, et al.* (1999) Human CD16 as a lysis receptor mediating direct natural killer cell cytotoxicity. *Proc Natl Acad Sci U S A* 96:5640-5644.

8.   Colucci F, Di Santo JP, & Leibson PJ (2002) Natural killer cell activation in mice and men: different triggers for similar weapons? *Nat Immunol* 3:807-813.

9.   Gomez-Roman VR*, et al.* (2006) A simplified method for the rapid fluorometric assessment of antibody-dependent cell-mediated cytotoxicity. *J Immunol Methods* 308:53-67.

10. Gomez-Roman VR*, et al.* (2005) Vaccine-elicited antibodies mediate antibody-dependent cellular cytotoxicity correlated with significantly reduced acute viremia in rhesus macaques challenged with SIVmac251. *J Immunol* 174:2185-2189.

11. Gupta N*, et al.* (2005) Targeted lysis of HIV-infected cells by natural killer cells armed and triggered by a recombinant immunoglobulin fusion protein: implications for immunotherapy. *Virology* 332:491-497.

12. Hessell AJ*, et al.* (2007) Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* 449:101-104.

13. Parren PW*, et al.* (2001) Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. *J Virol* 75:8340-8347.

14. Weiss CD & White JM (1993) Characterization of stable Chinese hamster ovary cells expressing wild-type, secreted, and glycosylphosphatidylinositol-anchored human immunodeficiency virus type 1 envelope glycoprotein. *J Virol* 67:7060-7066.

15. Yu H, Rabson AB, Kaul M, Ron Y, & Dougherty JP (1996) Inducible human immunodeficiency virus type 1 packaging cell lines. *J Virol* 70:4530-4537.

16. Cao J, Park IW, Cooper A, & Sodroski J (1996) Molecular determinants of acute single-cell lysis by human immunodeficiency virus type 1. *J Virol* 70:1340-1354.

17. Poignard P, Fouts T, Naniche D, Moore J, & Sattentau Q (1996) Neutralizing antibodies to human immunodeficiency virus type-1 gp120 induce envelope glycoprotein subunit dissociation. *J Exp Med* 183:473-484.

18. Frey G, *et al.* (2008) A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. *Proc Natl Acad Sci U S A* 105:3739-3744.

19. Gustchina E, Bewley CA, & Clore GM (2008) Sequestering of the prehairpin intermediate of gp41 by peptide N36Mut(e,g) potentiates the human immunodeficiency virus type 1 neutralizing activity of monoclonal antibodies directed against the N-terminal helical repeat of gp41. *J Virol* 82:10032-10041.

20. Sondermann P, Huber R, Oosthuizen V, & Jacob U (2000) The 3.2-A crystal structure of the human IgG1 Fc fragment-Fc gammaRIII complex. *Nature* 406:267-273.

21. Chesla SE, Li P, Nagarajan S, Selvaraj P, & Zhu C (2000) The membrane anchor influences ligand binding two-dimensional kinetic rates and three-dimensional affinity of FcgammaRIII (CD16). *J Biol Chem* 275:10235-10246.

22. Zanetti G, Briggs JA, Grunewald K, Sattentau QJ, & Fuller SD (2006) Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. *PLoS Pathog* 2(8):790-797.

23. Calarese DA, *et al.* (2003) Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 300:2065-2071.

24. Liu J, Bartesaghi A, Borgnia MJ, Sapiro G, & Subramaniam S (2008) Molecular architecture of native HIV-1 gp120 trimers. *Nature* 455:109-113.

25. Cardoso RM, *et al.* (2005) Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity* 22:163-173.

26. Ofek G*, et al.* (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.

27. Buchacher AR*, et al.* (1992) Human monoclonal antibodies against gp41 and gp120 as potential agent for passive immunization. *Vaccines* 92:191-195.

28. Saphire EO*, et al.* (2001) Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science* 293:1155-1159.

29. Egan MA, Carruth LM, Rowell JF, Yu X, & Siliciano RF (1996) Human immunodeficiency virus type 1 envelope protein endocytosis mediated by a highly conserved intrinsic internalization signal in the cytoplasmic domain of gp41 is suppressed in the presence of the Pr55gag precursor protein. *J Virol* 70:6547-6556.

30. Ye L*, et al.* (2004) Surface stability and immunogenicity of the human immunodeficiency virus envelope glycoprotein: role of the cytoplasmic domain. *J Virol* 78:13409-13419.

31. Trkola A*, et al.* (2005) Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat Med*. 11(6):615-622.

32. Lazar GA*, et al.* (2006) Engineered antibody Fc variants with enhanced effector function. *Proc Natl Acad Sci U S A* 103:4005-4010.

APPENDIX A

This appendix contains the software code written for two programs, *VictorExtract* and *GraphExtract* for use in Microsoft Excel 2004 with Visual Basic for Applications. The former was written to process the raw data output of a Victor3 luminometer (Perkin Elmer) for the neutralization assay and fit the data to a one-site dose-response model with fixed minima and maxima of 0 and 100% by minimization of the residual sum of squares. The later program processes the output from *VictorExtract* and arranges it in a format for creating figures using graphical software programs such as KaleidaGraph and Prism.

```
' Joshua S. Klein, California Institute of Technology, Copyright 2006

Option Explicit
Public filePath, infoArray(10, 14), resultsArray() As Variant
Public fileList, fileName, alphaStr, rssqStr As String
Public x, y, z, a, b, c, e, zz, yy, xx, aa, bb As Integer
Public plateArray(8, 12), d, neutValue, jacobianValue, totneutValue,
avgneutValue, rssqValue, chireducedValue As Double

Sub VictorExtract()
    UserForm1.Show
End Sub
Sub extractData()
Dim jacobianArray(1, 8)
UserForm1.Hide
'Read userform entries
For x = 1 To 10
    infoArray(x, 1) = UserForm1.Controls(x - 1).Value
    infoArray(x, 2) = UserForm1.Controls(x + 16).Value
    infoArray(x, 3) = UserForm1.Controls(x + 138).Value
    infoArray(x, 4) = UserForm1.Controls(x + 28).Value
    infoArray(x, 5) = UserForm1.Controls(x + 39).Value
    infoArray(x, 6) = UserForm1.Controls(x + 50).Value
    infoArray(x, 7) = UserForm1.Controls(x + 61).Value
    infoArray(x, 8) = UserForm1.Controls(x + 72).Value
    infoArray(x, 9) = UserForm1.Controls(x + 83).Value
    infoArray(x, 10) = UserForm1.Controls(x + 94).Value
    infoArray(x, 11) = UserForm1.Controls(x + 105).Value
    infoArray(x, 12) = UserForm1.Controls(x + 116).Value
    infoArray(x, 13) = UserForm1.Controls(x + 127).Value
Next x
'Assign plate columns per antibody
For x = 1 To 10
    If Trim(infoArray(x, 1)) <> "" Then
        For y = 4 To 13
            If infoArray(x, y) = "True" Then infoArray(x, 14) =
infoArray(x, 14) & Str(y - 1)
        Next y
    End If
Next x
'Select and open source data file
filePath = Application.GetOpenFilename()
MsgBox (filePath)
Application.Workbooks.Open (filePath)
x = 1
y = Len(filePath)
While x > 0
    If Mid(filePath, y, 1) <> ":" Then
        y = y - 1
        Else: x = 0
    End If
Wend
'Assign contents of Plate worksheet to array and close source data file
fileName = Right(filePath, Len(filePath) - y)
Application.Workbooks(fileName).Activate
Application.Worksheets("Plate").Activate
For x = 1 To 12
    For y = 1 To 8
```

```
        plateArray(y, x) = ActiveSheet.Cells(y + 6, x).Value
    Next y
Next x
Application.Workbooks(fileName).Close
'Create new data workbook
Application.Workbooks.Add
Application.ActiveSheet.Name = fileName
Application.ActiveSheet.Cells(1, 1).Value = "Date"
Application.ActiveSheet.Cells(1, 2).Value = "Antibody"
Application.ActiveSheet.Cells(1, 3).Value = "Isolate"
Application.ActiveSheet.Cells(1, 4).Value = "IC50 (nM)"
Application.ActiveSheet.Cells(1, 5).Value = "IC50 Err (nM)"
Application.ActiveSheet.Cells(1, 6).Value = "IC50 (ug/mL)"
Application.ActiveSheet.Cells(1, 7).Value = "IC50 Err (ug/mL)"
Application.ActiveSheet.Cells(1, 8).Value = "RSSQ"
Application.ActiveSheet.Cells(1, 9).Value = "Row A (nM)"
Application.ActiveSheet.Cells(1, 10).Value = "Row B (nM)"
Application.ActiveSheet.Cells(1, 11).Value = "Row C (nM)"
Application.ActiveSheet.Cells(1, 12).Value = "Row D (nM)"
Application.ActiveSheet.Cells(1, 13).Value = "Row E (nM)"
Application.ActiveSheet.Cells(1, 14).Value = "Row F (nM)"
Application.ActiveSheet.Cells(1, 15).Value = "Row G (nM)"
Application.ActiveSheet.Cells(1, 16).Value = "Row H (nM)"
Application.ActiveSheet.Cells(1, 17).Value = "GM1"
Application.ActiveSheet.Cells(1, 18).Value = "GM2"
Application.ActiveSheet.Cells(1, 19).Value = "GM3"
Application.ActiveSheet.Cells(1, 20).Value = "GM4"
Application.ActiveSheet.Cells(1, 21).Value = "GM5"
Application.ActiveSheet.Cells(1, 22).Value = "GM6"
Application.ActiveSheet.Cells(1, 23).Value = "GM7"
Application.ActiveSheet.Cells(1, 24).Value = "GM8"
Application.ActiveSheet.Cells(1, 25).Value = "Avg GM"
Application.ActiveSheet.Cells(1, 26).Value = "VC1"
Application.ActiveSheet.Cells(1, 27).Value = "VC2"
Application.ActiveSheet.Cells(1, 28).Value = "VC3"
Application.ActiveSheet.Cells(1, 29).Value = "VC4"
Application.ActiveSheet.Cells(1, 30).Value = "VC5"
Application.ActiveSheet.Cells(1, 31).Value = "VC6"
Application.ActiveSheet.Cells(1, 32).Value = "VC7"
Application.ActiveSheet.Cells(1, 33).Value = "VC8"
Application.ActiveSheet.Cells(1, 34).Value = "Avg VC"
alphaStr = "ABCDEFGHIJKLMNOPQRSTUVWXYZ"
b = 1
For y = 1 To 10
    xx = 0
    rssqStr = ""
    a = Len(infoArray(y, 14))
    c = 0
    For x = 1 To a
        If Trim(Mid(infoArray(y, 14), x, 1)) <> "" Then
            If Val(Mid(infoArray(y, 14), x, 2)) > 2 Then
                c = (Val(Mid(infoArray(y, 14), x, 2)))
                'MsgBox (c)
                Application.ActiveSheet.Cells(b + 1, 1).Value =
UserForm1.TextBox12.Value
                Application.ActiveSheet.Cells(b + 1, 2).Value =
infoArray(y, 1)
```

```
            Application.ActiveSheet.Cells(b + 1, 3).Value =
UserForm1.TextBox11.Value
                If Val(infoArray(y, 2)) > 0 Then
                    d = Val(infoArray(y, 2))
                    Else: d = 0
                End If
                For z = 9 To 16
                    Application.ActiveSheet.Cells(b + 1, z).Value = d
                    d = d / 3
                Next z
                'Set initial IC50 value to median concentration
                Application.ActiveSheet.Cells(b + 1, 4).Value =
Application.ActiveSheet.Cells(b + 1, 12).Value
                'Calculate average GM and average VC
                Application.ActiveSheet.Cells(b + 1, 25).Value =
"=average(Q" & b + 1 & ":X" & b + 1 & ")"
                Application.ActiveSheet.Cells(b + 1, 34).Value =
"=average(Z" & b + 1 & ":AG" & b + 1 & ")"
                'Enter RL values
                For z = 1 To 8
                    yy = 0
                    Application.ActiveSheet.Cells(b + 1, z + 16).Value
= plateArray(z, 1)
                    Application.ActiveSheet.Cells(b + 1, z + 25).Value
= plateArray(z, 2)
                    While yy < 1000
                        If Application.ActiveSheet.Cells(b + 1, z * 4 +
31 + yy).Value = "" Then
                            Application.ActiveSheet.Cells(b + 1, z * 4
+ 31 + yy).Value = plateArray(z, c)
                            yy = 1000
                            Else: yy = yy + 16
                        End If
                    Wend
                Next z
                'Calculate % neutralization and fitted values
                For z = 1 To 8
                    yy = 0
                    While Application.ActiveSheet.Cells(b + 1, z * 4 +
31 + yy).Value <> ""
                        If LCase(Application.ActiveSheet.Cells(b + 1, z
* 4 + 31 + yy).Value) <> "x" Then
                            'Application.ActiveSheet.Cells(b + 1, z
* 4 + 31 + yy + 1).Value = (1 - ((Application.ActiveSheet.Cells(b + 1,
z * 4 + 31 + yy).Value - _
                            Application.ActiveSheet.Cells(b +
1, 25).Value) / (Application.ActiveSheet.Cells(b + 1, 34).Value - _
                            Application.ActiveSheet.Cells(b +
1, 25).Value))) * 100
                            'Calculate obs % neutralization
                            Application.ActiveSheet.Cells(b + 1, z
* 4 + 31 + yy + 1).Value = "=(1 - (" & Application.ActiveSheet.Cells(b
+ 1, z * 4 + 31 + yy).Address(RowAbsolute:=False) & " - " & _
                            Application.ActiveSheet.Cells(b +
1, 25).Address(RowAbsolute:=False) & ")/(" &
Application.ActiveSheet.Cells(b + 1, 34).Address(RowAbsolute:=False) &
" - " & _
```

```
                                                Application.ActiveSheet.Cells(b +
1, 25).Address(RowAbsolute:=False) & ")) * 100"
                                        'Calculate expected % neutralization
                                        Application.ActiveSheet.Cells(b + 1, z
* 4 + 31 + yy + 2).Value = "=100/(1+D" & b + 1 & "/" & Mid(alphaStr, 17
- z, 1) & b + 1 & ")"
                                        'Calculate residual
                                        Application.ActiveSheet.Cells(b + 1, z
* 4 + 31 + yy + 3).Value = "=(" & Mid(Application.ActiveSheet.Cells(b +
1, z * 4 + 31 + yy _
                                        + 1).Address(RowAbsolute:=False),
2) & "-" & Mid(Application.ActiveSheet.Cells(b + 1, z * 4 + 31 + yy +
2).Address(RowAbsolute:=False), 2) & ")^2"
                                    End If
                                    yy = yy + 16
                                    xx = xx + 1
                        Wend
                    Next z
                End If
            c = 0
            End If
        Next x
        'Calculate RSSQ = sum(y_obs - y_fit)^2, n = 1 to i
        If xx > 0 Then
            Application.ActiveSheet.Cells(b + 1, ((yy - 16) / 32 + 1) * 32
+ 35).Value = "end"
            For z = 1 To 8
                rssqStr = rssqStr & "(" &
Mid(Application.ActiveSheet.Cells(b + 1, z * 4 +
33).Address(RowAbsolute:=False), 2) & " - average("
                For x = 1 To xx / 16
                    'Ignore cells with values of "x" (assumes manual
deletion)
                    If LCase(Application.ActiveSheet.Cells(b + 1, z * 4 +
31 + x * 32 - 32).Value) <> "x" Then
                        rssqStr = rssqStr &
Mid(Application.ActiveSheet.Cells(b + 1, z * 4 + 32 + x * 32 -
32).Address(RowAbsolute:=False), 2) & ","
                    End If
                Next x
                rssqStr = Left(rssqStr, Len(rssqStr) - 1) & "))^2,"
            Next z
            ActiveSheet.Cells(b + 1, 8).Value = "=sum(" & Left(rssqStr,
Len(rssqStr) - 1) & ")"
        End If
        b = b + 1
Next y
'Minimize RSSQ using built-in solver function
For x = 1 To 10
    If Application.ActiveSheet.Cells(x + 1, 8).Value > 0 Then
        SolverReset
        SolverOK
setCell:=Application.ActiveSheet.Cells.Range(ActiveSheet.Cells(x + 1,
8), ActiveSheet.Cells(x + 1, 8)), _
        maxMinVal:=2, _

byChange:=Application.ActiveSheet.Cells.Range(ActiveSheet.Cells(x + 1,
4), ActiveSheet.Cells(x + 1, 4))
```

```
        SolverSolve UserFinish:=True
        'Calculate error
        xx = 1
        While LCase(ActiveSheet.Cells(x + 1, xx + 34).Value) <> "end"
            xx = xx + 1
        Wend
        xx = (xx - 1) / 32
        'Calculate asymptotic standard error
        jacobianValue = 0
        totneutValue = 0
        avgneutValue = 0
        rssqValue = 0
        chireducedValue = 0
        For aa = 1 To 8
            'Jacobian = d(% neut)/d(IC50)
            jacobianValue = jacobianValue + (((-100 /
ActiveSheet.Cells(x + 1, aa + 8).Value) / (1 + ActiveSheet.Cells(x + 1,
4).Value / _
                ActiveSheet.Cells(x + 1, aa + 8).Value) ^ 2)) ^ 2
        Next aa
        'For aa = 1 To 8
        '    If xx > 0 Then
        '        For bb = 1 To xx
        '            'Ignore deleted values
        '            If LCase(ActiveSheet.Cells(x + 1, bb * 32 + aa * 4
- 4).Value) <> "x" Then
        '                totneutValue = totneutValue +
ActiveSheet.Cells(x + 1, bb * 32 + aa * 4).Value
        '            End If
        '        Next bb
        '        avgneutValue = totneutValue / xx
        '        Else:
        '            If LCase(ActiveSheet.Cells(x + 1, bb * 32 + aa * 4
- 4).Value) <> "x" Then
        '                avgneutValue = ActiveSheet.Cells(x + 1, bb *
32 + aa * 4).Value
        '            End If
        '    End If
        '    If avgneutValue <> 0 Then
        '        rssqValue = rssqValue + (avgneutValue -
ActiveSheet.Cells(x + 1, aa * 4 + 37 - 4).Value) ^ 2
        '    End If
        '    totneutValue = 0
        'Next aa
        ActiveSheet.Cells(x + 1, 5).Value = (jacobianValue ^ -0.5) *
((Application.ActiveSheet.Cells(x + 1, 8).Value / 7) ^ 0.5)
        If Application.ActiveSheet.Cells(x + 1, 4).Value > 0 Then
            'Calculate IC50  and error in ug/mL
            If Val(infoArray(x, 3)) > 0 Then
                Application.ActiveSheet.Cells(x + 1, 6).Value =
Application.ActiveSheet.Cells(x + 1, 4).Value * (10 ^ -9) *
infoArray(x, 3) * (10 ^ 3)
                Application.ActiveSheet.Cells(x + 1, 7).Value =
Application.ActiveSheet.Cells(x + 1, 5).Value * (10 ^ -9) *
infoArray(x, 3) * (10 ^ 3)
            End If
        End If
    End If
```

```
Next x
'Format results
Rows("1:1").Select
Selection.Font.Bold = True
Rows("1:10").Select
Selection.Columns.AutoFit
Selection.Columns.HorizontalAlignment = xlCenter
End Sub
 Sub GraphExtract()
    UserForm2.Show
 End Sub
 Sub getCurves()
    UserForm2.Hide
    alphaStr = "ABCDEFGHIJKLMNOPQRSTUVWXYZ"
    xx = 1
    While LCase(ActiveSheet.Cells((UserForm2.TextBox1.Value), xx +
34).Value) <> "end"
        xx = xx + 1
    Wend
    xx = (xx - 1) / 32
    ReDim resultsArray(3, 8 + xx, 8) As Variant
    resultsArray(1, 1, 1) = "Date"
    resultsArray(1, 2, 1) = "Antibody"
    resultsArray(1, 3, 1) = "Isolate"
    resultsArray(1, 4, 1) = "GM"
    resultsArray(1, 5, 1) = "VC"
    resultsArray(1, 6, 1) = "Conc (nM)"
    resultsArray(1, 7, 1) = "Avg % neut"
    resultsArray(1, 8, 1) = "Err % neut"
    resultsArray(2, 1, 1) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), 1).Value
    resultsArray(2, 2, 1) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), 2).Value
    resultsArray(2, 3, 1) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), 3).Value
    For x = 1 To 8
        resultsArray(2, 4, x) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), x + 16).Value
        resultsArray(2, 5, x) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), x + 25).Value
        resultsArray(2, 6, x) =
Application.ActiveSheet.Cells((UserForm2.TextBox1.Value), 17 - x).Value
    Next x
    For aa = 1 To 8
        'Get neutralization values for each replicate
        For bb = 1 To xx
            resultsArray(1, 8 + bb, 1) = "Exp " & bb
            If Val(ActiveSheet.Cells((UserForm2.TextBox1.Value), bb *
32 + aa * 4 - 1).Value) <> 0 Then
                resultsArray(2, 8 + bb, aa) =
ActiveSheet.Cells((UserForm2.TextBox1.Value), bb * 32 + aa * 4).Value
                resultsArray(2, 7, aa) = resultsArray(2, 7, aa) +
resultsArray(2, 8 + bb, aa)
                Else: resultsArray(2, 8 + bb, aa) = "x"
            End If
        Next bb
        'Calculate avg and stdev neut
        If xx > 1 Then
```

```
            resultsArray(2, 7, aa) = "=average(I" & aa + 1 & ":" &
Mid(alphaStr, xx + 8, 1) & aa + 1 & ")"
            resultsArray(2, 8, aa) = "=stdev(I" & aa + 1 & ":" &
Mid(alphaStr, xx + 8, 1) & aa + 1 & ")"
        End If
    Next aa
    Application.Workbooks.Add
    For a = 1 To UBound(resultsArray, 2)
        Application.ActiveSheet.Cells(1, a).Value = resultsArray(1, a,
1)
        For x = 1 To 8
            Application.ActiveSheet.Cells(x + 1, a).Value =
resultsArray(2, a, x)
        Next x
    Next a
    Application.ActiveSheet.Cells(10, 3).Value = "Average"
    Application.ActiveSheet.Cells(11, 3).Value = "St dev"
    Application.ActiveSheet.Cells(10, 4).Value = "=average(D2:D9)"
    Application.ActiveSheet.Cells(10, 5).Value = "=average(E2:E9)"
    Application.ActiveSheet.Cells(11, 4).Value = "=stdev(D2:D9)"
    Application.ActiveSheet.Cells(11, 5).Value = "=stdev(E2:E9)"
    'Format results
    Rows("1:1").Select
    Selection.Font.Bold = True
    Rows("10:11").Select
    Selection.Font.Bold = True
    Rows("1:11").Select
    Selection.Columns.AutoFit
    Selection.Columns.HorizontalAlignment = xlCenter
End Sub
```

Victor Extract user form:

Graph Extract user form:

APPENDIX B

This appendix lists the 93 different alternative antibody architecture and Fc mutant constructs that were cloned, expressed, and tested for neutralization where expression was successful (N.D., not done).

	Name	Notes	Cloned	Expresses	Neutralizes
1	(CD4-GCN4)$_3$	Trimeric GCN4 design	√	X	
2	(CD4-GCN4)$_4$	Tetrametic GCN4 design	√	X	
3	(CD4-GCN4)$_6$	Hexameric GCN4 design	√	X	
4	(CD4-scFvE51)$_3$	Trimeric GCN4 design	√	X	
5	(CD4-scFvE51)$_4$	Tetrametic GCN4 design	√	X	
6	(dsFvb12-GCN4)$_3$-VLG44C/VHQ105C	Trimeric GCN4 design/disulfide Fv	√	X	
7	(dsFvb12-GCN4)$_4$-VLG44C/VHQ105C	Tetrametic GCN4 design/disulfide Fv	√	X	
8	(Fab 2G12-GCN4)$_4$	Tetrametic GCN4 design	√	X	
9	(Fab 2G12-GCN4)$_6$	Hexameric GCN4 design	√	X	
10	(Fab b12-GCN4)$_3$	Trimeric GCN4 design	√	X	
11	(Fab b12-GCN4)$_4$	Tetrametic GCN4 design	√	X	
12	Bs(scFv)$_2$-Fab 4E10-2F5		√	X	
13	Bs(scFv)$_4$-IgG 4E10-2F5		√	X	
14	diabody 2F5-4E10		√	√	√
15	diabody 4E10-4E10		√	√	√
16	diabody b12-4E10		√	√	√
17	diabody b12-b12		√	√	√
18	IgG 2F5 (G$_4$S)$_1$	Random coil extension	√	√	√
19	IgG 2F5 (G$_4$S)$_2$	Random coil extension	√	N.D.	
20	IgG 2F5 (G$_4$S)$_3$	Random coil extension	√	√	√
21	IgG 2F5 (G$_4$S)$_4$	Random coil extension	√	N.D.	
22	IgG 2F5 (G$_4$S)$_5$	Random coil extension	√	√	√
23	IgG 2F5 (G$_4$S)$_6$	Random coil extension	√	N.D.	
24	IgG 2F5 (G$_4$S)$_7$	Random coil extension	√	X	
25	IgG 2F5 CD4	Protein fusion extension	√	X	
26	IgG 2F5 invasin	Protein fusion extension	√	X	
27	IgG 2F5 invasin$_{1-3}$	Protein fusion extension	√	√	√
28	IgG 2F5 neogenin	Protein fusion extension	√	X	
29	IgG 2F5 S239D	ADCC enhancement	√	√	N.D.
30	IgG 2F5 S239D/I332E	ADCC enhancement	√	√	N.D.
31	IgG 2F5 titin$_{168-170}$	Protein fusion extension	√	X	
32	IgG 2G12 (G$_4$S)$_1$	Random coil extension	√	√	√
33	IgG 2G12 (G$_4$S)$_3$	Random coil extension	√	√	√
34	IgG 2G12 (G$_4$S)$_5$	Random coil extension	√	√	√
35	IgG 2G12 (G$_4$S)$_7$	Random coil extension	√	√	√
36	IgG 2G12 CD4	Protein fusion extension	√	X	
37	IgG 2G12 invasin	Protein fusion extension	√	√	√
38	IgG 2G12 invasin1-3	Protein fusion extension	√	√	√
39	IgG 2G12 neogenin	Protein fusion extension	√	√	√
40	IgG 2G12 S239D	ADCC enhancement	√	√	N.D.
41	IgG 2G12 S239D/I332E	ADCC enhancement	√	√	N.D.
42	IgG 2G12 titin$_{168-170}$	Protein fusion extension	√	X	
43	IgG 4E10 (G$_4$S)$_1$	Random coil extension	√	√	√
44	IgG 4E10 (G$_4$S)$_3$	Random coil extension	√	√	√
45	IgG 4E10 (G$_4$S)$_5$	Random coil extension	√	√	√
46	IgG 4E10 (G$_4$S)$_7$	Random coil extension	√	X	
47	IgG 4E10 CD4	Protein fusion extension	√	X	

48	IgG 4E10 invasin	Protein fusion extension	√	X	
49	IgG 4E10 invasin$_{1-3}$	Protein fusion extension	√	√	√
50	IgG 4E10 neogenin	Protein fusion extension	√	√	√
51	IgG 4E10 S239D	ADCC enhancement	√	√	N.D.
52	IgG 4E10 S239D/I332E	ADCC enhancement	√	√	N.D.
53	IgG scFv4E10	scFv fusion instead of Fab	√	√	√
54	IgG 4E10 titin 168-170	Protein fusion extension	√	X	
55	IgG b12 (G$_4$S)$_1$	Random coil extension	√	√	√
56	IgG b12 (G$_4$S)$_2$	Random coil extension	√	N.D.	
57	IgG b12 (G$_4$S)$_3$	Random coil extension	√	√	√
58	IgG b12 (G$_4$S)$_4$	Random coil extension	√	N.D.	
59	IgG b12 (G$_4$S)$_5$	Random coil extension	√	√	√
60	IgG b12 (G$_4$S)$_6$	Random coil extension	√	N.D.	
61	IgG b12 (G$_4$S)$_7$	Random coil extension	√	X	
62	IgG b12 CD4	Protein fusion extension	√	X	
63	IgG b12 invasin	Protein fusion extension	√	X	
64	IgG b12 invasin$_{1-3}$	Protein fusion extension	√	√	√
65	IgG b12 neogenin	Protein fusion extension	√	X	
66	IgG b12 S239D	ADCC enhancement	√	√	N.D.
67	IgG b12 S239D/I332E	ADCC enhancement	√	√	N.D.
68	IgG b12 scFv	scFv fusion instead of Fab	√	√	√
69	IgG b12 titin$_{168-170}$	Protein fusion extension	√	X	
70	IgG dsFv4E10-(G$_4$S)$_2$-VLG44C/VHQ105C	disulfide Fv fusion	√	√	√
71	IgG dsFv4E10-(G$_4$S)$_7$-VLG44C/VHQ105C	disulfide Fv fusion	√	√	√
72	IgG dsFvb12-(G$_4$S)$_2$-VLG44C/VHQ105C	disulfide Fv fusion	√	N.D.	
73	IgG dsFvb12-(G$_4$S)$_2$-VLG44C/VHQ105C	disulfide Fv fusion	√	N.D.	
74	IgG GFP	To make heterodimers	√	√	X
75	IgG sCD4-17b		√	√	√
76	IgG T1249	C-peptide	√	√	√
77	IgG T20	C-peptide	√	√	√
78	IgG T2635	C-peptide	√	√	√
79	IgG T651	C-peptide	√	√	√
80	scBvFv 2F5-4E10		√	√	√
81	scBvFv 4E10-4E10		√	√	√
82	scBvFv b12-4E10		√	√	√
83	scBvFv b12-b12		√	√	√
84	sCD4		√	√	√
85	sCD4-17b		√	√	√
86	scFv 2F5		√	√	√
87	scFv 2G12		√	N.D.	
88	scFv 4E10		√	√	√
89	scFv b12		√	√	√
90	scFv4E10-TACE-Fc-scFvb12	Random coil extension	√	√	√
91	scFv4E10-TACE-Fc-scFvb12	Random coil extension	√	√	√
92	Tet(scFv)$_4$ 4E10-2F5-b12-17b		√	X	
93	Tri(scFv)$_6$ 4E10-2F5-b12		√	X	

APPENDIX C

My contribution to this work was in the characterization of the binding properties of non-glycosylated human transferrin mutants to the soluble extracellular portion of human transferrin receptor by surface plasmon resonance.

123

Composition of pH-Sensitive Triad in C-Lobe of Human Serum Transferrin. Comparison to Sequences of Ovotransferrin and Lactoferrin Provides Insight into Functional Differences in Iron Release[†]

Peter J. Halbrooks,[‡] Anthony M. Giannetti,[§,⊥] Joshua S. Klein,[§] Pamela J. Björkman,[○] Julia R. Larouche,[‡]
Valerie C. Smith,[||] Ross T. A. MacGillivray,[||] Stephen J. Everse,[‡] and Anne B. Mason*,[‡]

*Department of Biochemistry, University of Vermont, College of Medicine, Burlington, Vermont 05405, Graduate Option in
Biochemistry and Molecular Biophysics, Department of Biology, and Howard Hughes Medical Institute, California Institute of
Technology, Pasadena, California 91125, and Department of Biochemistry and Molecular Biology, University of British
Columbia, Vancouver, British Columbia V6T 1Z3, Canada*

Received September 14, 2005; Revised Manuscript Received September 29, 2005

ABSTRACT: The transferrins (TF) are a family of bilobal glycoproteins that tightly bind ferric iron. Each
of the homologous N- and C-lobes contains a single iron-binding site situated in a deep cleft. Human
serum transferrin (hTF) serves as the iron transport protein in the blood; circulating transferrin binds to
receptors on the cell surface, and the complex is internalized by endocytosis. Within the cell, a reduction
in pH leads to iron release from hTF in a receptor-dependent process resulting in a large conformational
change in each lobe. In the hTF N-lobe, two critical lysines facilitate this pH-dependent conformational
change allowing entry of a chelator to capture the iron. In the C-lobe, the lysine pair is replaced by a triad
of residues: Lys534, Arg632, and Asp634. Previous studies show that mutation of any of these triad
residues to alanine results in significant retardation of iron release at both pH 7.4 and pH 5.6. In the
present work, the role of the three residues is probed further by conversion to the residues observed at the
equivalent positions in ovotransferrin (Q-K-L) and human lactoferrin (K-N-N) as well as a triad with an
interchanged lysine and arginine (K534R/R632K). As expected, all of the constructs bind iron and associate
with the receptor with nearly the same K_D as the wild-type monoferric hTF control. However, interesting
differences in the effect of the substitutions on the iron release rate in the presence and absence of the
receptor at pH 5.6 are observed. Additionally, titration with KCl indicates that position 632 must have a
positively charged residue to elicit a robust rate acceleration as a function of increasing salt. On the basis
of these observations, a model for iron release from the hTF C-lobe is proposed. These studies provide
insight into the importance of charge and geometry of the amino acids at these positions as a partial
explanation for differences in behavior of individual TF family members, human serum transferrin,
ovotransferrin, and lactoferrin. The studies collectively highlight important features common to both the
N- and C-lobes of TF and the critical role of the receptor in iron release.

The transferrins (TF)[1] comprise a family of iron-binding
proteins that include serum transferrin, ovotransferrin (oTF),
and lactoferrin (LTF) (*1*). The primary function of serum
transferrin synthesized by the liver and secreted into the blood
plasma is iron transport. Human TF (hTF) binds Fe^{3+} and
delivers it to cells via a receptor-mediated endocytotic

pathway (*2, 3*). Ovotransferrin, synthesized in the oviduct
and comprising ~12% of avian egg whites, serves as an anti-
microbial agent by chelating any available Fe^{3+}, thus,
depriving invading bacteria of this metal which is essential
to their proliferation (*4*). Interestingly, the identical protein
(differing only in the composition of the single carbohydrate)

* Correspondence should be addressed to: Dr. Anne B. Mason,
Department of Biochemistry, B402 Given Building, 89 Beaumont
Avenue, Burlington, VT 05405. Phone, 802-656-0343; fax, 802-862-
8229; e-mail, anne.mason@uvm.edu.
‡ University of Vermont.
§ Graduate Option in Biochemistry and Molecular Biophysics,
California Institute of Technology.
⊥ Current address, Roche Palo Alto, LLC, Palo Alto, California
94043.
○ Department of Biology and Howard Hughes Medical Institute,
California Institute of Technology.
|| University of British Columbia.

[1] Abbreviations: TF, transferrin or transferrins; hTF, human serum
transferrin; oTF, ovotransferrin; LTF, lactoferrin; N-His hTF-NG,
recombinant nonglycosylated human serum transferrin with a Factor
Xa cleavage site and a hexa-His tag attached to the amino-terminus of
the protein; N-His Y95F/Y188F hTF-NG, monoferric hTF with iron
in the C-lobe; N-His Y426F/Y517F hTF-NG, monoferric hTF with iron
in the N-lobe; N-His Y95F/Y188F/Y426F/Y517F hTF-NG, apo-hTF;
TFR, transferrin receptor 1; sTFR, soluble transferrin receptor 1;
DMEM-F12, Dulbecco's modified Eagle's medium-Ham F-12 nutrient
mixture; BHK cells, baby hamster kidney cells; UG, Ultroser G; FBS,
fetal bovine serum; BSA, bovine serum albumin; HRP, horseradish
peroxidase; TMB, 3,3′,5,5′-tetramethylbenzidine; BA, butyric acid;
Tiron, 4,5-dihydroxy-1, 3-benzenedisulfonic acid; EDTA, ethylenedi-
amine-tetracetic acid; HEPES, *N*-(2-hydroxyethyl) piperazine-*N*′-
ethanesulfonic acid; MES, morpholinoethanesulfonic acid; Ni-NTA,
nickel-nitrilotriacetic acid; SPR, surface plasmon resonance; RU,
resonance units; KISAB, kinetically significant anion-binding.

Halbrooks et al.

also transports iron in avian plasma (5), thereby fulfilling two roles in a tissue specific manner. LTF, found in milk, tears, saliva, and other mammalian secretions, functions as a powerful antimicrobial agent in these secretions (6).

Almost all transferrins are composed of two homologous lobes designated the amino-terminal lobe (N-lobe) and the carboxyl-terminal lobe (C-lobe). The lobes share ~40% sequence identity and appear to have arisen from at least one gene duplication and fusion event (7). The iron-liganding residues in each lobe of oTF, LTF, and hTF are identical and are composed of an aspartic acid, two tyrosines, and a histidine. The two other ligands that make up the distorted octahedral geometry are provided by a synergistic carbonate anion, which is bound in a concomitant manner with iron (2). The iron is held in a cleft comprised of two subdomains (termed NI and NII, and CI and CII). Each lobe has the ability to bind Fe^{3+} reversibly in a pH-dependent manner such that iron release is accelerated at acidic pH and slowed at basic pH; factors influencing the mechanism of iron release within the cell include receptor-binding, temperature, and an as yet unidentified chelator (8−10).

Although the identical amino acids serve as ligands to the iron, there are significant differences in the iron-binding and release properties of the N- and C-lobes of any given TF and also between TF family members (1, 11−13). For example, our recent studies document that at pH 5.6, under identical conditions, the rate of release from the C-lobe is at least 200-fold slower than the release rate from the N-lobe (14). The observed functional dissimilarity between lobes has been attributed to differences in the hydrogen bonding network within the cleft and the "second shell" residues (defined as amino acids that share hydrogen bonds with the iron-binding ligands) (11, 15, 16).

Analysis of the crystal structure of the hen oTF N-lobe led Dewan et al. (12) to propose that a pH-sensitive dilysine pair, made up of a lysine residue from domain NI and one from domain NII, might play a role in eliciting the conformational change. In the iron-containing hTF N-lobe, Lys206 is 3.04 Å from Lys296, but 9 Å apart in the apo-structure (17, 18). In fact, this dilysine pair has been shown to play a key role in the mechanism leading to cleft opening and iron release (19, 20). A structural feature of the two lysine residues is the presence of a "hydrophobic box" (composed of the two liganding tyrosines, Tyr95 and Tyr188, as well as Tyr85 and the histidine ligand, His249), which almost completely surrounds them. As shown in Table 1, these residues are highly conserved. This hydrophobic environment apparently changes the pK_a of one or both the lysine residues such that they are able to share a single proton at neutral pH (19). Mathematical and modeling studies of the N-lobe by Rinaldo and Field (21) emphasize the importance of the two lysines in iron release but suggest that the mechanism involves transfer of a proton from Lys296 to the nearby Tyr188 ligand which weakens the hold on iron and results in cleft opening.

The C-lobe of hTF has a triad of residues (Lys534, Arg632, and Asp634) predicted by Dewan et al. (12) to be important in the mechanism leading to iron release. Our studies have demonstrated that mutation of any of the triad residues to alanine completely inhibited iron release at pH 7.4 and slowed the rate of release substantially at pH 5.6 (13). Interestingly, in oTF and LTF, these three residues are not

Table 1: Alignments of Critical Residues in the N- and C-Lobes of Various TFs[a]

human TF	rabbit TF	pig TF	LTF	oTF
		N-lobe		
Glu83	Glu83	Glu82	Glu80	Glu80
Tyr85	Tyr85	Tyr84	Tyr82	Tyr82
Tyr95	**Tyr95**	**Tyr94**	**Tyr92**	**Tyr92**
Tyr188	**Tyr188**	**Tyr192**	**Tyr192**	**Tyr191**
Lys206	**Lys206**	**Lys210**	*Arg210*	**Lys209**
His249	**His249**	**His253**	**His253**	**His250**
Lys296	Lys296	Lys300	Lys301	Lys301
		C-lobe		
Glu410	Glu410	Glu414	Glu413	Glu413
Tyr412	Tyr412	Tyr416	Tyr415	Ty415
Tyr426	**Tyr425**	**Tyr431**	**Tyr435**	**Tyr431**
Tyr517	**Tyr514**	**Tyr526**	**Tyr528**	**Tyr524**
Lys534	**Lys531**	**Lys543**	**Lys546**	*Gln541*
His585	**His582**	**His594**	**His597**	**His592**
Arg632	Arg629	Arg641	*Asn644*	*Lys638*
Asp634	Asp631	Asp643	*Asn646*	*Leu640*

[a] The residues in bold designate the NII and CII subdomains in each lobe.

conserved (see Table 1), leading to the hypothesis that some of the altered iron release properties displayed by the C-lobes of these proteins might be due to these differences. The only studies in which iron release from these three family members has been directly compared are chemical relaxation studies by El Hage Chahine et al. (22−25). In this work, acid-induced iron release was measured and showed that hTF and oTF have proton-assisted carbonate loss from the N-lobe as the first step. Release of iron from the C-lobe seems to follow a pathway that is more complex than found for the N-lobe. Although specific differences were noted in the release rates, in the precise number of protons taken up, and in the rate-limiting steps from the C-lobes of oTF and hTF, according to these studies, they are mechanistically similar (22, 24). In contrast, these chemical relaxation studies showed that the steps leading to iron release from LTF differ mechanistically from hTF and oTF (23).

Binding of hTF to the specific transferrin receptor 1 (TFR) has a substantial effect on the rate of iron release, although the precise steps leading to receptor-mediated iron release from hTF are not completely clear (26). A number of studies have identified putative residues in both TFR and TF involved in their interaction (27−29). This work shows that specific residues within the CI domain of the C-lobe are critical to high affinity binding of hTF to the TFR. Additionally, our recent collaborative work clearly indicates that mutation of His349 in the CI domain to alanine or to a lysine residue (as found in oTF) totally eliminates the acceleration of iron release by the soluble TFR in vitro (30). These results demonstrate that the interaction of His349 with the TFR is essential to receptor-mediated release of the iron ~30 Å distant from this histidine and further imply an induced conformational change which leads to a stabilization of the apo-conformation.

In the current study, conversion of the K534-R632-D634 residues in hTF to the residues found at the same positions in oTF (Q-K-L) and human LTF (K-N-N) has been undertaken to help determine how the composition of the three residues influences the rate of iron release from the various C-lobes in the presence and absence of the TFR. Additionally, a K534R/R632K switch mutant was designed to alter

the geometry, without changing the charge of the two residues. In each of the full-length constructs, the two tyrosine ligands (Y95 and Y188) were mutated to phenylalanine to eliminate iron binding in the N-lobe (*31*). The ability of each recombinant monoferric hTF to bind iron, associate with the soluble TFR, and undergo pH-mediated iron release in both the presence and absence of the TFR is reported. Additionally, the effect of salt on iron release has been examined to determine whether any of the three residues might serve as an anion-binding site (*32*, *33*). Evaluation of the properties of each mutant contributes to a better understanding of the intrinsic differences between the three family members from which they are derived.

MATERIALS AND METHODS

Materials. Dulbecco's modified Eagle's medium-Ham F-12 nutrient mixture (DMEM-F12), and antibiotic-antimycotic solution (100×) were from Gibco-BRL-Life Technologies. Fetal bovine serum (FBS) was obtained from Atlanta Biologicals (Norcross, GA). Ultroser G (UG) is a serum replacement from BioSepra (Cergy, France). Bovine factor Xa was purchased from Haematologic Tecnologies, Inc. (Essex Junction, VT). The Quik-change mutagenesis kit was obtained from Strategene. Ni-NTA resin came from Qiagen. Corning expanded surface roller bottles and Dynatech Immunolon 4 Removawells were obtained from Fisher Scientific. Hi-Prep 26/60 Sephacryl S-200HR and S300HR columns were from Amersham Pharmacia. Methotrexate from Bedford Laboratories was purchased at a local hospital pharmacy and used for selection of plasmid-containing cells. Centricon 30 microconcentrators, YM-30 ultrafiltration membranes, and a spiral cartridge concentrator (CH2PRS) fitted with an S1Y10 cartridge were all from Millipore/Amicon. Bovine serum albumin (BSA) was from Sigma. Rabbit anti-mouse immunoglobulin G was from Southern Biological Associates. Immunopure NHS-LC-Biotin and immunopure avidin-horseradish peroxidase were from Pierce, as was the India His probe-HRP. The TMB Microwell peroxidase substrate system was obtained from Kirkegaard and Perry Laboratories (Gaithersburg, MD). All other chemicals and reagents were of analytical grade.

Plasmid Generation of hTF Mutants. The Quik-change protocol from Stratagene was used to introduce mutations into the hTF cDNA in the pNUT vector as described previously (*31*, *34*). All of the hTF constructs contained an N-terminal hexa-His tag and lacked N-linked glycosylation sites due to the presence of the two mutations, N413D and N611D (*31*, *34*, *35*). Additionally, each mutated triad was introduced into full-length hTF containing the Y95F/Y188F double mutant that lacks the ability to bind iron in the N-lobe (*31*). The complete DNA sequence of all clones was determined on both strands prior to transfection of the plasmid into BHK cells. As shown below, complementary mutagenic oligonucleotide primers containing the following mutations were used to introduce the desired mutations. The mutagenic nucleotides are shown in bold, underlined type:

K534Q, 5′ GAT GTG GCC TTT GTG **C**AA CAC CAG ACT GTC CC 3′; K534R, 5′ GGA GAT GTG GCC TTT GTG **AG**A CAC CAG ACT GTC 3′; R632K/D634L, 5′GAA ACC **AA**G GAC CTT CTG TTC A**AA** GAT **CT**C ACA GTA TGT TTG GCC 3′; and R632N/D634N, 5′GAA ACC

AAG GAC CTT CTG TTC A**AT** GAT **AA**T ACA GTA TGT TTG GCC 3′.

Production of the single point mutants in which each member of the triad has been converted to alanine has been described in detail previously (*13*). Native ovotransferrin was purified from chicken egg white (*36*).

Production and Purification of Recombinant Proteins. Expression and purification of the recombinant hTFs was carried out as previously described (*31*). Briefly, the recombinant TFs were secreted from BHK cells into the surrounding tissue culture medium. The recombinant hTF was isolated from the media using Qiagen Ni-NTA and Sephacryl S200HR columns to yield a single band on Coomassie blue stained SDS−PAGE gels. The hTF was recovered from the S200HR column in 100 mM NH_4HCO_3, concentrated to 15 mg/mL, and stored at −20 °C until use. The amount of recombinant hTF in the tissue culture media and column eluates prior to the final step was determined by a competitive immunoassay (*37*). Following the gel filtration column, the concentration was determined by spectral analysis using extinction coefficients as previously reported (*31*). The recombinant His-tagged sTFR from the baculovirus insect cell system was expressed and purified by a similar protocol, as described in detail previously (*38*).

Removal of the Hexa-His Tag. Removal of the His tag was required for hTF samples used for the surface plasmon resonance-based binding assay in a Biacore instrument (see below) because His-tagged samples, while tightly binding to sTFR, exhibited nonconcentration-dependent response levels that could not be directly interpreted. His tag removal followed the protocol described in detail previously (*14*).

Spectral Data. To assess iron binding for each hTF, UV−vis spectra were recorded on a Varian Cary 100 spectrophotometer (600−250 nm). Difference spectra were generated by subtracting the spectrum of HEPES buffer (50 mM, pH 7.4) from the sample spectra. The reported results are the average values derived from at least three such scans.

Affinity Measurements Using Surface Plasmon Resonance. A BIACORE 2000 biosensor system (Amersham Biosciences) was used to assay the interaction between sTFR and TF as described previously (*29*, *39*). Binding of injected hTF to sTFR immobilized on the sensor chip results in changes in surface plasmon resonance that are read out in real time as resonance units (RU) (*40*, *41*). Purified sTFR was coupled to the flow cells of a CM5 sensor chip (Amersham Biosciences) using random amine coupling as previously described (*38*). Wild-type or mutant hTF was injected over the flow-cells at 70 μL/min at 25 °C in 50 mM PIPES buffer, pH 7.5, 150 mM NaCl, and 0.005% P20 surfactant. Raw sensorgram data was preprocessed using the Scrubber software package (Biologic Software Pty.; http://www.biologic.com.au) and globally fit to 2:1 or 1:1 models in Clamp 99. Results from 2:1 models were statistically corrected for better comparison with the results of the 1:1 model as previously described (*30*, *42*).

Production of hTF/sTFR Complex. A molar excess of the control N-His Y95F/Y188F hTF NG (or hTF mutants) in 100 mM NH_4HCO_3 was incubated with sTFR for 30 min at room temperature. Following passage through a 13 mm Millex-GV syringe filter (0.2 μm), each sample was loaded onto a HiPrep 26/60 Sephacryl S300HR gel filtration column (Amersham Biosciences) to separate the hTF/sTFR complex

from the free hTF. The column was equilibrated and run in 100 mM NH$_4$HCO$_3$ at a flow rate of 1−1.5 mL/min using a BioCad Sprint chromatography system, and baseline separation was achieved between the complex and the hTF. The identity of the hTF/TFR complex and the free hTF was verified by SDS−PAGE analysis on 10% gels.

Kinetics of Iron Release. As described previously in detail (*10*), the kinetics of iron removal from hTF at pH 5.6 were measured in 100 mM MES buffer using 4 mM EDTA as the chelating agent. The buffer−chelator mixture was allowed to equilibrate prior to addition of the hTF samples, and release of iron was monitored directly by following the decrease in absorbance at 293 nm on a dual-beam Cary 100 spectrophotometer equipped with a thermostable Peltier block accessory. The hTF was present at a final concentration of ∼6 μM, and all experiments were conducted at 25 °C. To ensure complete iron removal, data was collected until at least three half-lives had elapsed.

In assays involving TF/sTFR complexes, lower concentrations (∼500 nM) of protein ensured receptor solubility at the acidic pH of 5.6. A QuantaMaster Spectrofluorometer (Photon Technology International, South Brunswick NJ) was used to measure fluorescence output. This instrument is equipped with a 75-W Xenon arc lamp as an excitation source and excitation/emission monochromators. Fluorescence emission spectra were measured by exciting the sample at 280 nm and collecting the emission at 330 nm. Experiments were carried out at 25 °C with an entry slit of 0.25 nm and an exit slit of 1.0 nm. Data were recorded at 1 s intervals immediately following the addition of either the complex with the sTFR (yielding a final hTF concentration of 500 nM) or hTF (also 500 nM) alone to a magnetically stirred cuvette containing 1.8 mL at a final concentration of 100 mM MES, pH 5.6, and 4 mM EDTA. As described in the Results, the KCl concentration varied from 150 to 500 mM. It is noteworthy that attempts to measure iron release from the hTF/sTFR complex in the absence of added salt were unsuccessful, presumably due to the tendency of the sTFR to precipitate from solution at pH 5.6 in the absence of salt.

RESULTS

Spectral Analysis. To reduce the complexity and allow comparison of the properties of recombinant hTFs with mutations in the C-lobe triad, all of the constructs contain mutated tyrosine ligands that have been shown previously to eliminate iron binding in the N-lobe (*31*). This strategy simplifies the interpretation of the results, since in the absence of iron binding in the N-lobe any observed iron-binding and iron-release properties can be completely attributed to the C-lobe. Spectral scans of each of the recombinant proteins were taken, since intrinsic spectral parameters (λ_{max} and A_{280}) clearly indicate whether the recombinant proteins are able to bind iron and whether the mutations result in a change in the visible maximum. As shown in Table 2, all of the recombinant proteins had a visible absorbance maximum consistent with specific binding of iron in the C-lobe. The slight upfield shift of λ_{max} in the oTF and LTF triads and downfield shift for the K534R/R632K switch reflect subtle changes in the geometry of the iron-binding site, in particular the relationship of iron to the two tyrosine ligands. We note

Table 2: Spectral Scans of Recombinant hTF Proteins

protein	λ_{max} (nm)	A_{280}/A_{max}	note
N-His hTF-NG[a]	464	27.5	diferric
N-His K206E hTF-NG[a]	463	25.6	diferric
N-His Y95F/Y188F hTF-NG[b]	461	42.2	apo-N
N-His Y95F/Y188F/**R632N/ D634N** hTF-NG	467	42.2	LTF triad
N-His Y95F/Y188F/**K534R/ R632K** hTF-NG	453	44.5	switch
N-His Y95F/Y188F/**K534Q/ R632K/D634L** hTF-NG	463	42.8	oTF triad

[a] Reported previously (*13*). [b] Reported previously (*31*).

Table 3: Comparison of Binding Affinities of the Triad Mutants for sTFR at pH 7.4 As Determined by Surface Plasmon Resonance

sample	K_{D1} (nM)	K_{D2} (nM)
hTF-NG	1.84	31.8
Y95F/Y188F hTF-NG	28.5	
Y95F/Y188F "oTF" triad hTF-NG	24.2	
Y95F/Y188F "LTF" triad hTF-NG	39.4	
Y95F/Y188F "Switch" hTF-NG	25.2	

that, since the triad residues are part of the hydrogen-binding network holding the two subdomains of the C-lobe together, they clearly could have an effect on the visible spectrum.

Surface Plasmon Resonance. A surface plasmon resonance (SPR)-based binding assay was used to evaluate the binding affinity of each triad mutant to the soluble portion of the human TFR. Two controls were performed, one with native hTF and one with the Y95F/Y188F hTF NG (monoferric C-lobe). On the basis of the K_D values, obtained for fitting the sensorgram data, the monoferric C-lobe species exhibits a lower affinity for sTFR (Table 3), clearly confirming the preference of sTFR for diferric hTF. Interestingly, a 1:1 binding model, which yields a single K_D value, was sufficient to model the monoferric C-lobe data, whereas a 2:1 bivalent ligand model, which yields two K_D values, is necessary to fit the curves for diferric hTF control (as previously reported) (*29*, *30*). At pH 7.4, the monoferric control has a 15-fold reduced binding affinity compared to K_{D1} of diferric hTF (but bound with an affinity which is similar to K_{D2}). The three triad mutants range from a 0.8- to 1.4-fold change in K_D relative to the monoferric C-lobe control, demonstrating that the composition of the residues at positions 534−632−634 does not substantially alter the strength of the hTF/sTFR interaction in vitro.

Kinetics of Iron Release as a Function of Salt Concentration. Iron-release rates from salt titrations of the hTF proteins containing the human, oTF, LTF, and K534R/R632K triads at pH 5.6 and 25 °C are presented in Figure 1. The mutant containing the three residues found in oTF releases iron marginally faster than does the control N-His Y95F/Y188F hTF-NG and also demonstrates the same increase in the rate constant as a function of increasing KCl. As shown in the inset of Figure 1, when the three residues found in LTF are present, iron release is dramatically slowed and the salt effect is also significantly muted. Additionally, iron release is slowed by the presence of the K534R/R632K switch, although the salt effect is retained.

Kinetic assays were performed on each triad mutant free or bound to sTFR at pH 5.6 in the presence of 300 mM KCl. This pH was chosen because it is the putative
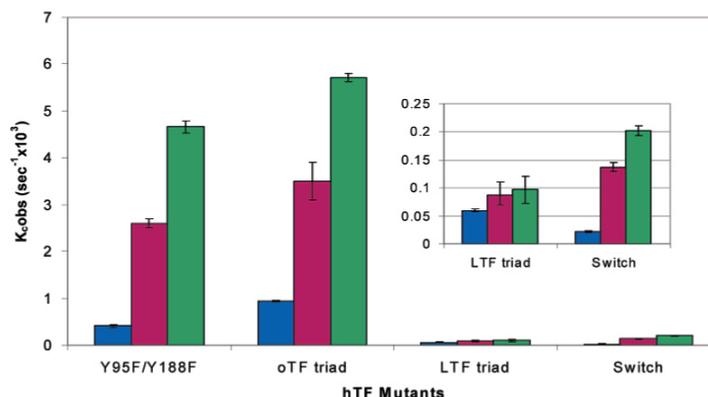
FIGURE 1: Iron-release kinetics for hTF species titrated with KCl. Iron-release rates from the monoferric C-lobe control (N-His Y95F/Y188F hTF-NG) and each of the triad mutants in 100 mM MES, pH 5.6, with 4 mM EDTA as the chelator at 25 °C. Three salt concentrations were used, 0 M (blue), 0.3 M (red), and 0.5 M (green) KCl. The inset expands the *Y*-axis to more clearly show the effect of salt on the iron release rate of LTF and the switch.

Table 4: Kinetics of Iron Release from TFs with and without sTFR[a]

protein[b]		$k_{obs} \pm SD$[c] ($s^{-1} \times 10^{-3}$)	n	fold difference TFR/no TFR
"hTF" triad (K-R-D)		2.6 ± 0.2[d]	4	
	+ TFR	273.0 ± 45[d]	4	105
"oTF" triad (Q-K-L)		3.3 ± 0.1	3	
	+ TFR	91.3 ± 4.7	3	27
authentic oTF C-lobe		1.5 ± 0.002	2	
"LTF" triad (K-N-N)		0.087 ± 0.017	5	
	+ TFR	25.5 ± 0.9	3	195
"switch" triad (R-K-D)		0.136 ± 0.006	4	
	+ TFR	16.5 ± 0.1	3	124
K534A		0.018 ± 0.001	2	
	+ TFR	6.7 ± 0.3	3	364
R632A		too slow		
	+ TFR	2.3 ± 0.1	4	
D634A		too slow		
	+ TFR	too slow		

[a] Iron release from 500 nM hTF at pH 5.6 (100 mM MES, 4 mM EDTA, and 300 mM KCl) in the absence or presence of the soluble TFR at 25 °C. The single point alanine scanning mutants are included for comparisons as is the kinetic rate from the C-lobe of native oTF, all determined under identical conditions. [b] The control is N-HisY95F/Y188F hTF-NG. All constructs are in this background. Note that the triad in hTF is K534/R632/D634. [c] Measurements were carried out as described in Materials and Methods on either a PTI (Photon Technology International QuantaMaster) spectrofluorometer or a Varian Cary 100 dual beam spectrophotometer. [d] Previously reported (*30*).

physiological pH at which receptor-mediated iron release is believed to occur (*3, 43*). The results (Table 4) show that in all cases binding of the mutants to sTFR dramatically increases the rate of iron release, as observed in previous studies (*26, 30*). Of the mutants, the oTF "triad" was least affected by the presence of receptor (∼27-fold increase in release rate relative to oTF "triad" alone), while the rate of release from wild-type triad, the LTF "triad", and the K534R/R632K switch was accelerated 100-fold or more when TFR was present. Additionally, for comparison, the release rates from the three mutants in which each triad residue is converted to an alanine are included in Table 4. As previously

reported (*13*), in each case, this conversion results in a much slower rate of iron release. In fact, under the stated conditions, we were unable to measure any release from the D634A mutant either in the presence or absence of the TFR, and although we were able to obtain a release rate for the other two alanine mutants (K534A and R634A) in the presence of the TFR, the release of iron from the R632A mutant was too slow to measure in its absence.

We have found, as previously reported (*44*), that the release rates measured for selected samples on the PTI and the Cary 100 yield very similar iron release rates regardless of whether they were fluorescence- or absorbance-based (data not shown). From an experimental standpoint, an absorbance-based approach is more convenient over an extended period of time (for example, the LTF and switch triad mutants in the absence of sTFR), because multiple samples can be run simultaneously. Alternatively, the fluorescence-based approach requires a lower concentration of protein, critical for the measurement of the hTF/sTFR complex where solubility is limiting.

Since the recombinant hTF containing the three residues found in LTF released iron so slowly, a salt titration was performed with the complex to determine the release rates in a more experimentally tractable time frame. As shown in Table 5, a marginal salt effect is observed in both the presence and absence of sTFR. In either case (with or without the sTFR), the differences in the observed release rates at the various KCl concentrations are very small compared to the increases observed in the other mutants (Figure 1).

DISCUSSION

A compelling goal of transferrin research is to provide a structural explanation for the well-known differences between family members and between the two lobes of each family member with respect to their thermodynamic and kinetic properties. Over a period of years, detailed studies of iron uptake and iron release from hTF, oTF, and LTF have been undertaken by El Hage Chahine and his colleagues with the aim of elucidating the mechanistic differences between each

Halbrooks et al.

Table 5: Salt Titration of LTF-Triad Containing hTF with and without sTFR[a]

protein	[KCl] (mM)	$K_{obs} \pm$ SD ($s^{-1} \times 10^3$)	n
R632N/D634N "LTF"	0	0.060 ± 0.002	5
	150	0.113 ± 0.006	3
	300	0.087 ± 0.019	5
	500	0.097 ± 0.027	5
R632N/D634N "LTF"+ TFR	0	could not do	-
	150	19.3 ±0.6	3
	300	25.5 ± 0.9	3
	500	31.6 ±1.6	3

[a] Rate of iron release from the N-His Y95F/Y188F LTF triad in 100 mM MES, pH 5.6, and 4 mM EDTA, at 25 °C as a function of salt concentration.

family member (*22*, *23*, *25*, *45*, *46*). As mentioned in the Introduction, many interesting insights were obtained from this work. The challenge is now to relate the mechanistic details to specific residues within each lobe of each family member. A triad of residues in the C-lobe of hTF has clearly been shown to be important in the mechanism of iron release (*12*, *13*). The key contribution of each member of the triad in iron release is confirmed and extended by the kinetic data for the triad alanine mutants shown in Table 4. Under the stated conditions, the rate of release in the absence of the TFR was too slow to measure for the R632A and D634A mutants and 144-fold slower than the control for the K534A mutant. The current studies were carried out to specifically evaluate the effect of substitution of these three residues with residues found at the equivalent positions in family members, oTF and LTF.

Information on the iron-containing human C-lobe is currently limited to an unrefined 3.3 Å resolution structure (*47*). Fortunately, structures of the closely related pig TF (2.15 Å) and rabbit TF (2.6 Å) are available (*48*), providing a snapshot of the orientation of the three residues in the iron-containing C-lobe. As shown in Table 1, there is a complete conservation of the triad residues in human, pig, and rabbit TF. As depicted in Figure 2, residues Arg632 and Asp634 are in the CI domain, with Lys534 opposite them in the CII domain. For clarity, in the following discussion, human numbering will be used (although the residue numbers for each species are provided in Table 1). The pig structure reveals that there are two potential hydrogen bonds between the aspartic acid and the arginine (both in the CI domain) and a single hydrogen bond between the aspartic acid and the lysine in the CII domain. In the pig structure, the NZ group of the lysine is 4.09 Å from the NE group of the arginine. The lower resolution rabbit structure shows a very similar pattern, although in this case, the lysine NZ group and arginine NE groups are closer (3.3 Å). Additionally, the triad residues reside in a hydrophobic box composed of the three iron-ligands (Tyr426, Tyr517, and His585) as well as the highly conserved Tyr at position 412. As a result, cation-π interactions responsible for the unusual pK_a values noted in the dilysine pair in the N-lobe probably occur here also. In the N-lobe, the ability of the two lysines, Lys206 and Lys296, to share a hydrogen bond is directly attributed to the abnormal pK_a values created by the hydrophobic environment. Likewise, the pK_a values of the residues comprising the triad very likely differ considerably from "normal" values (*19*, *21*).
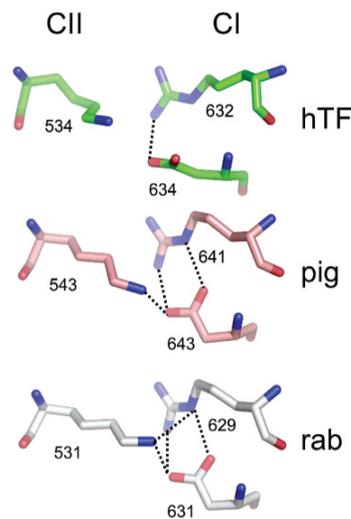
FIGURE 2: Triad residues in TF family members. The diferric structures were obtained from the PDB (1H76 for pig serum TF and 1JNF for rabbit serum TF). Potential H-bonds and residue positions are indicated.

As briefly mentioned in the Introduction, at least two different (but related) pathways have been suggested to describe iron release from the N-lobe (*12*, *21*). Given the considerable sequence similarity between lobes, certain aspects of these pathways also undoubtedly apply to the C-lobe. The decrease in pH within the acidic endosome results in an influx of protons which would be predicted to disrupt the hydrogen bonds between the aspartic acid and the arginine and lysine residues of the hTF triad. The positively charged lysine and arginine residues on opposite sides of the cleft might then repel each other to initiate the opening of the cleft, as suggested for the two lysines in the N-lobe (*12*). Alternatively, Rinaldo and Field (*21*) have proposed a concerted proton transfer such that Arg632 could acquire a proton from either Lys534 (or from a water molecule that is within hydrogen-bonding distance of Lys534 in the pig structure) and subsequently transfer a proton to the nearby tyrosine ligand (Tyr517) disrupting its hold on the iron. According to the model offered by Rinaldo and Field, this transfer is an early step in a concerted release mechanism leading to the break up of the hydrogen-bond network within the cleft and allowing entry of a chelator. In both the N- and C-lobe of hTF, this mechanism also appears to involve the histidine ligand which is held in position by interaction with a highly conserved glutamate residue (Glu83 in the N-lobe and Glu410 in the C-lobe, see Table 1). Since Arg632 in the C-lobe has a pK_a that is two units higher than the corresponding lysine (Lys296) in the N-lobe, protonation of the tyrosine ligand would presumably require a lower pH than needed for the N-lobe with its dilysine pair. Similarly, if a water molecule is involved in the process of proton transfer, a more extreme pH would also be required (*21*). Although experimental evidence strongly supports the lower pH requirement for iron removal from the C-lobe (*22–24*,

49), distinguishing experimentally between these two mechanisms is difficult. The 4.09 Å distance between the lysine and arginine residues in the C-lobe of the pig structure might appear to favor the second pathway, although reorientation of the Lys–Arg pair in response to anion binding, the change in pH, and/or binding to TFR also may be possible, especially considering the well-known flexibility within the cleft (*48*). In either case, the result is cleft loosening or opening and access of the chelator to the iron within the cleft.

Surprisingly, substitution of the residues found in oTF for the hTF triad only marginally altered the iron-release rate from the C-lobe (Table 4 and Figure 1), despite the lack of conservation of any residues at the equivalent positions (Table 1). The crystal structure of diferric oTF (PDB 1OVT) shows that the OE1 group of the glutamine (equivalent to Lys534) is 2.96 Å from the NZ group of the lysine (equivalent to Arg632), thereby possibly participating in a hydrogen bond linking the CI and CII subdomains (Table 1). For native oTF, the slower rate of iron release from the C-lobe of oTF compared to the N-lobe has been attributed to the difference between the Gln–Lys and Lys–Lys pairs in each lobe, respectively (*50, 51*). As described above, for the mammalian TFs, these residues in each lobe of oTF exist in a hydrophobic environment comprised of the liganding residues and the highly conserved tyrosine residue (Tyr85 in the N-lobe and Tyr412 in the C-lobe, Table 1) (*52*). In the present study, determination of the rate of iron release for the C-lobe of native oTF under our stated conditions yielded a rate that is approximately half of that measured for the oTF triad in the hTF background (Table 4). Obviously, it is not clear whether placement of the glutamine and lysine residues within the human scaffolding would allow them to interact, since glutamine is shorter than arginine. The absence of the hydrogen bond might explain the faster rate in the oTF "triad" mutant compared to native oTF.

As discussed above, proton transfer from a lysine residue to tyrosine is probably more facile than transfer from the arginine residue found in hTF and might at least partially account for the slightly faster iron release from the oTF "triad" compared to hTF. In the native oTF structure, the leucine residue at the position equivalent to Asp634 has no contact with the lysine or the glutamine residues, thereby invalidating the triad designation. The absence of the negatively charged aspartic acid residue interacting with both the lysine and arginine residues in hTF implies that the pH dependence of iron release from the C-lobe of oTF should be different, although the chemical relaxation experiments do not seem to support this contention (*23*).

To further evaluate the effect of substitution of lysine by glutamine, the iron release rates for the K206Q and K296Q mutants in the N-lobe were examined. The crystal structure (1.8 Å) of the K206Q mutant (*53*) shows that the Gln206 OE1 side chain is 2.9 Å from the NZ of Lys296. Significantly, the rates of iron release from these two mutants are *not* equivalent (450-fold slower for K206Q versus 130-fold slower for K296Q, compared to the wild-type N-lobe), demonstrating the importance of geometry and orientation of the two residues. The rate of release from the N-lobe (in the absence of salt) is 83.0×10^{-3} s^{-1} at pH 5.6 versus 0.183 $\times 10^{-3}$ s^{-1} for the K206Q mutant and 0.633×10^{-3} s^{-1} for the K296Q mutant (*8*). The release rates for these two N-lobe mutants appear to fall into a range with the same order of

magnitude as the rates measured for the C-lobe of the hTF and the oTF "triad" (0.383×10^{-3} s^{-1} and 0.933×10^{-3} s^{-1}, respectively, in the absence of salt).

In contrast to the amino acid substitutions made to simulate the residues in oTF, the substitutions made to reproduce the LTF triad (K-N-N) dramatically altered the iron-release kinetics, resulting in a 30-fold slower rate of iron release compared to hTF (Table 4 and Figure 1). As shown in Table 1, the residues in the LTF triad vary at the middle and last positions, with two asparagine residues in place of arginine and aspartic acid, respectively. Significantly, this K-N-N motif is conserved in the C-lobe of the seven LTF sequences published to date (*54*). Examination of the 2.2 Å diferric human LTF structure (PDB 1FCK) reveals that the carbonyl oxygen of the lysine residue shares a hydrogen bond with a threonine residue (equivalent to Thr537 in hTF) (*55, 56*). The ND2 group of the asparagine residue at the position equivalent to Arg632 in hTF is in potential hydrogen-bonding distance of the highly conserved Glu410 and the ND1 of the His585 ligand. The asparagine at the position equivalent to Asp634 in hTF does not appear to be in contact with any other residue. Since in the native LTF structure none of the members of the "triad" are in hydrogen-bonding distance with each other, they, like the oTF triad, do not actually comprise a "triad" per se. As mentioned previously, in the N-lobe of hTF, Glu83 (which occupies the equivalent position to Glu410) interacts with the histidine ligand and appears to be involved in the concerted mechanism of iron release (*16, 21*), a relationship which may be retained in hTF containing the LTF substitutions. However, the apparent absence of interactions between the three residues might be predicted to retard iron release from the LTF "triad" mutant as is observed; there appears to be no possibility of any kind of "triggering" event either by transfer of a proton or repulsion of residues on opposite sides of the cleft.

The K534R/R632K switch also significantly slows iron release from the C-lobe (Figure 2 and Table 4). It is noteworthy that interchanging the identities of two residues that are both positively charged results in altered iron-release properties as well as a 10 nm shift in the absorbance maximum (Table 2). Evidently, the geometry of these residues is just as important as the charge in the mechanism of iron release, reminiscent of work describing the R210K mutation in LTF N-lobe (*57, 58*) as well as the K206Q and K296Q N-lobe hTF mutants discussed above. It is reasonable that the K534R/R632K switch leads to similar subtle shifts in the neighboring side chains, resulting in the altered functional behavior. Obviously, a structure is needed to determine whether the aspartic acid still interacts with the lysine and arginine residues in their reversed positions, although the slowness in the release rate argues against this.

Complexes with the soluble human TFR and each of the "triad" mutants were prepared, and the rate of iron release was determined. As a first step, it was necessary to show whether introduction of other residues for those comprising the hTF triad had any effect on the ability of the resulting mutants to bind to the TFR. As expected, no significant differences in binding of any of the "triad" mutants to the soluble human TFR were found in our SPR experiments (Table 3). Consistent with previous reports, the monoferric C-lobe control (Y95F/Y188F hTF-NG) exhibited a lower

Halbrooks et al.

affinity (15-fold weaker binding) to sTFR than the diferric control (native hTF-NG) (*59−61*). Additionally, the mutants with the oTF, LTF, and K534R/R632K conversions bound to sTFR with K_D values similar to the K_D of the control monoferric parent protein. The results are completely consistent with previous studies of the hTF/sTFR interaction in which none of the residues comprising the triad have been implicated in TFR binding (*27, 28*). Nevertheless, given the dynamic nature of the interaction of hTF with the sTFR, we felt that this needed to be verified experimentally.

In agreement with previous work, the presence of the sTFR significantly accelerates the rate of iron release from the hTF C-lobe at pH 5.6 (*26, 30*); thus, the rate of iron release from the native human C-lobe of hTF increases at least 100-fold due to the presence of the sTFR. This TFR-induced acceleration is attributed to stabilization of the apo or open/partially open form of hTF which is induced by interaction of His349 with a hydrophobic patch on the TFR (*30*). This work strongly supports earlier suggestions of long-range-induced conformational changes with impact on the iron site. Interestingly, the effect of the TFR on the rate of iron release is considerably less for the oTF triad. The rate of release induced by binding of the oTF "triad" mutant to the sTFR is ∼4-fold lower than the increase observed for the hTF C-lobe in the sTFR complex. This is a significant difference and implies that this "triad" mutant does not achieve the full conformational change required for maximal stimulation of iron release from the C-lobe. It is tempting to attribute some of the difference to the missing aspartic acid residue.

Perhaps most interesting is the LTF mutant (K-N-N) which seems to lack any kind of mechanism to open the cleft and is also missing a robust salt effect (see below), resulting in a 30- fold decrease in the rate of iron release in the absence of the TFR (in comparison to hTF). In the presence of the TFR, the rate of release from the hTF with the LTF triad is increased 230-fold compared to the ∼100-fold enhancement of the rate of release from the C-lobe in the hTF/sTFR complex. Nevertheless, the actual rate is still ∼10-fold slower (0.026 vs 0.273 s^{-1}) (Table 4). Likewise, the rate of iron release from the switch mutant is 120-fold faster when bound to the TFR, although, again, the actual rate is the slowest of any of the triad mutants in complex with the TFR (0.0165 s^{-1}), some 17-fold-slower than the control. Thus, the composition of the triad residues has a very significant effect on the TFR-induced iron release from the C-lobe, reinforcing the idea that the residues in the cleft immediately surrounding the iron site communicate locally but also have a more global response. The extremely slow rate of release observed for the single alanine mutants even in the presence of the TFR (Table 4) confirms the roles of both charge and geometry in the process of iron release. The importance of the aspartic acid residue is highlighted by the fact that no iron release is observed either with or without the TFR in an experimentally feasible time frame.

A further goal of the triad substitutions was the possibility of identifying an anion-binding site in the C-lobe. It has long been known that anionic salts influence chelator-mediated iron removal from hTF (*8, 10*). Several types of anion interactions have been described, including kinetically significant anions, which are thought to bind at a site or sites near the metal-binding center and change the conformation allowing an incoming chelator to more readily remove the

iron (*33*). In the C-lobe, mutation of the lysine at position 569 to glutamine produced a 15-fold decrease in the iron-release rate compared to the control (*62*), indicating that this residue might be a kinetically significant anion binding site (or KISAB site). However, it was also suggested that Lys569 is probably not the only such site in the C-lobe. This conclusion appears to be supported by the fact that in pig TF the lysine at position 569 is replaced by a glutamic acid residue. The lack of conservation in this closely related family member implies that this residue is not essential to iron release (although it clearly has an effect). In the N-lobe, we have shown that the second shell residue network, specifically, the Lys296 and the synergistic anion anchoring Arg124, serve as anion-binding sites (*8, 10, 20*). The homologous residues in the C-lobe could similarly play such a role. Kinetic rates from the C-lobes containing the human, oTF, and the K534R/R632K switch "triads" all displayed a significant salt effect in which 6−9.5-fold increases between 0 and 500 mM KCl were observed (Figure 1). In all three constructs, the middle position has a positively charged residue (either a lysine or arginine). Interestingly, the mutant with the LTF substitutions at the triad positions displayed almost no dependence on KCl. To determine whether this mutant had a diminished salt effect or whether the slow iron-release kinetics was masking salt dependence, a salt titration was performed with the mutant/TFR complex (Table 4). Although, in the presence of the TFR, iron release is much more rapid, the effect of increasing the KCl concentration from 150 to 500 mM resulted in less than a 2-fold increase in the rate of iron release. Taken as a whole, the results suggest that the residue in the middle position might well be an anion site, since conversion of this residue to the nonbasic amino acid asparagine (in the LTF mutant) curtails the acceleration of the rate of release by salt. Somewhat surprisingly, we were unable to find any studies examining the effect of salt on iron release from LTF. The absence of a significant salt effect in the LTF highlights yet another fundamental difference to provide an explanation for the slower rate of iron release from native LTF in comparison to hTF. LTF seems ideally suited for its function of sequestering iron and withholding it from invading pathogens.

With the exception of the D634A mutant, none of the substitutions to the hTF triad completely eliminated the effect of the TFR in accelerating iron release, as might be predicted by the fact that all of the mutants retain His349 shown to be critical in this acceleration at acidic pH (*30*). Although neither native oTF nor native LTF bind to the human TFR (*63*), substitution of the three residues which comprise the triad in hTF did not have a significant effect on binding (as was expected). Clearly, in the absence of crystal structures, we do not know the exact relationship of the triad residues when placed in the human scaffolding. Nevertheless, the behaviors of the oTF and LTF mutants are instructive with regard to the importance of the identity of each member of the triad and of the geometry of the site. We note that the K-R-D triad combination in human TF seems particularly well-adapted to acceleration of iron release at pH 5.6 in the presence of the TFR, yielding by far the fastest rate. The aspartic acid residue provides a pH-sensitive element. As mentioned repeatedly, the key to understanding how the dilysine pair in the N-lobe and the triad in the C-lobe exist

is recognizing the hydrophobic environment in which these residues reside. This environment gives rise to unusual pK_a values, which allow anomalous behaviors, increasing or decreasing susceptibility to the protonation events which drive iron release.

REFERENCES

1. Baker, E. N., Baker, H. M., and Kidd, R. D. (2002) Lactoferrin and transferrin: functional variations on a common structural framework, *Biochem. Cell Biol. 80*, 27−34.
2. Aisen, P., Enns, C., and Wessling-Resnick, M. (2001) Chemistry and biology of eukaryotic iron metabolism, *Int. J. Biochem. Cell Biol. 33*, 940−959.
3. Klausner, R. D., Ashwell, G., van Renswoude, J., Harford, J. B., and Bridges, K. R. (1983) Binding of apotransferrin to K562 cells: explanation of the transferrin cycle, *Proc. Natl. Acad. Sci. U.S.A. 80*, 2263−2266.
4. Weinberg, E. D. (1974) Iron and susceptibility to infectious disease, *Science 184*, 952−956.
5. Williams, J. (1968) Comparison of glycopeptides from the ovotransferrin and serum transferrin of the hen, *Biochem. J. 108*, 57−67.
6. Lönnerdal, B., and Iyer, S. (1995) Lactoferrin: molecular structure and biological function, *Annu. Rev. Nutr. 15*, 93−110.
7. Williams, J. (1982) The evolution of transferrin, *Trends Biochem. Sci. 7*, 394−397.
8. He, Q.-Y., and Mason, A. B. (2002) Molecular aspects of release of iron from transferrins, in *Molecular and Cellular Iron Transport* (Templeton, D. M., Ed.) pp 95−123, Marcel Dekker, Inc., New York.
9. Harris, W. R., Cafferty, A. M., Abdollahi, S., and Trankler, K. (1998) Binding of monovalent anions to human serum transferrin, *Biochim. Biophys. Acta 1383*, 197−210.
10. He, Q. Y., Mason, A. B., Nguyen, V., MacGillivray, R. T. A., and Woodworth, R. C. (2000) The chloride effect is related to anion binding in determining the rate of iron release from the human transferrin N-lobe, *Biochem. J. 350*, 909−915.
11. Evans, R. W., Crawley, J. B., Joannou, C. L., and Sharma, N. D. (1999) Iron proteins, in *Iron and Infection: Molecular, Physiological and Clinical Aspects* (Bullen, J. J., and Griffiths, E., Eds.) 2nd ed., pp 27−86, John Wiley and Sons, Chichester, U.K.
12. Dewan, J. C., Mikami, B., Hirose, M., and Sacchettini, J. C. (1993) Structural evidence for a pH-sensitive dilysine trigger in the hen ovotransferrin N-lobe: implications for transferrin iron release, *Biochemistry 32*, 11963−11968.
13. Halbrooks, P. J., He, Q. Y., Briggs, S. K., Everse, S. J., Smith, V. C., MacGillivray, R. T. A., and Mason, A. B. (2003) Investigation of the mechanism of iron release from the C-lobe of human serum transferrin: mutational analysis of the role of a pH sensitive triad, *Biochemistry 42*, 3701−3707.
14. Mason, A. B., Halbrooks, P. J., James, N. G., Connolly, S. A., Larouche, J. R., Smith, V. C., MacGillivray, R. T. A., and Chasteen, N. D. (2005) Mutational analysis of C-lobe ligands of human serum transferrin: insights into the mechanism of iron release, *Biochemistry 44*, 8013−8021.
15. Baker, H. M., Anderson, B. F., Brodie, A. M., Shongwe, M. S., Smith, C. A., and Baker, E. N. (1996) Anion binding by transferrins: importance of second-shell effects revealed by the crystal structure of oxalate-substituted diferric lactoferrin, *Biochemistry 35*, 9007−9013.
16. He, Q.-Y., Mason, A. B., Woodworth, R. C., Tam, B. M., MacGillivray, R. T. A., Grady, J. K., and Chasteen, N. D. (1998) Mutations at nonliganding residues Tyr-85 and Glu-83 in the N-lobe of human serum transferrin−functional second shell effects, *J. Biol. Chem. 273*, 17018−17024.
17. MacGillivray, R. T. A., Moore, S. A., Chen, J., Anderson, B. F., Baker, H., Luo, Y. G., Bewley, M., Smith, C. A., Murphy, M. E., Wang, Y., Mason, A. B., Woodworth, R. C., Brayer, G. D., and Baker, E. N. (1998) Two high-resolution crystal structures of the recombinant N-lobe of human transferrin reveal a structural change implicated in iron release, *Biochemistry 37*, 7919−7928.
18. Jeffrey, P. D., Bewley, M. C., MacGillivray, R. T. A., Mason, A. B., Woodworth, R. C., and Baker, E. N. (1998) Ligand-induced conformational change in transferrins: crystal structure of the open form of the N-terminal half-molecule of human transferrin, *Biochemistry 37*, 13978−13986.
19. Steinlein, L. M., Ligman, C. M., Kessler, S., and Ikeda, R. A. (1998) Iron release is reduced by mutations of lysine 206 and 296 in recombinant N-terminal half-transferrin, *Biochemistry 37*, 13696−13703.
20. He, Q.-Y., Mason, A. B., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1999) Dual role of Lys206−Lys296 interaction in human transferrin N-lobe: iron-release trigger and anion-binding site, *Biochemistry 38*, 9704−9711.
21. Rinaldo, D., and Field, M. J. (2003) A computational study of the open and closed forms of the N-lobe human serum transferrin apoprotein, *Biophys. J. 85*, 3485−3501.
22. Abdallah, F. B., and El Hage Chahine, J. M. (1999) Transferrins, the mechanism of iron release by ovotransferrin, *Eur. J. Biochem. 263*, 912−920.
23. Abdallah, F. B., and El Hage Chahine, J. M. (2000) Transferrins: iron release from lactoferrin, *J. Mol. Biol. 303*, 255−266.
24. El Hage Chahine, J. M., and Pakdaman, R. (1995) Transferrin, a mechanism for iron release, *Eur. J. Biochem. 230*, 1102−1110.
25. El Hage Chahine, J.-M., and Fain, D. (1994) The mechanism of iron release from transferrin. Slow-proton-transfer-induced loss of nitrilotriacetatoiron(III) complex in acidic media, *Eur. J. Biochem. 223*, 581−587.
26. Zak, O., and Aisen, P. (2003) Iron release from transferrin, its C-lobe, and their complexes with transferrin receptor: presence of N-lobe accelerates release from C-lobe at endosomal pH, *Biochemistry 42*, 12330−12334.
27. Liu, R. T., Guan, J. Q., Zak, O., Aisen, P., and Chance, M. R. (2003) Structural reorganization of the transferrin C-lobe and transferrin receptor upon complex formation: the C-lobe binds to the receptor helical domain, *Biochemistry 42*, 12447−12454.
28. Cheng, Y., Zak, O., Aisen, P., Harrison, S. C., and Walz, T. (2004) Structure of the human transferrin receptor−transferrin complex, *Cell 116*, 565−576.
29. Giannetti, A. M., Snow, P. M., Zak, O., and Björkman, P. J. (2003) Mechanism for multiple ligand recognition by the human transferrin receptor, *PLoS Biol. 1*, 341−350.
30. Giannetti, A. M., Halbrooks, P. J., Mason, A. B., Vogt, T. M., Enns, C. A., and Björkman, P. J. (2005) The molecular mechanism for receptor-stimulated iron release from the plasma iron transport protein transferrin, *Structure*, in press.
31. Mason, A. B., Halbrooks, P. J., Larouche, J. R., Briggs, S. K., Moffett, M. L., Ramsey, J. E., Connolly, S. A., Smith, V. C., and MacGillivray, R. T. A. (2004) Expression, purification, and characterization of authentic monoferric and apo-human serum transferrins, *Protein Expression Purif. 36*, 318−326.
32. Marques, H. M., Watson, D. L., and Egan, T. J. (1991) Kinetics of iron removal from human serum monoferric transferrins by citrate, *Inorg. Chem. 30*, 3758−3762.
33. Egan, T. J., Ross, D. C., Purves, L. R., and Adams, P. A. (1992) Mechanism of iron release from human serum C-terminal mono-ferric transferrin to pyrophosphate: kinetic discrimination between alternative mechanisms, *Inorg. Chem. 31*, 1994−1998.
34. Mason, A. B., He, Q. Y., Halbrooks, P. J., Everse, S. J., Gumerov, D. R., Kaltashov, I. A., Smith, V. C., Hewitt, J., and MacGillivray, R. T. A. (2002) Differential effect of a His tag at the N- and C-termini: functional studies with recombinant human serum transferrin, *Biochemistry 41*, 9448−9454.
35. Mason, A. B., Miller, M. K., Funk, W. D., Banfield, D. K., Savage, K. J., Oliver, R. W. A., Green, B. N., MacGillivray, R. T. A., and Woodworth, R. C. (1993) Expression of glycosylated and nonglycosylated human transferrin in mammalian cells. Characterization of the recombinant proteins with comparison to three commercially available transferrins, *Biochemistry 32*, 5472−5479.
36. Williams, S. C., and Woodworth, R. C. (1973) The interaction of iron−conalbumin (anion) complexes with chick embryo red blood cells, *J. Biol. Chem. 248*, 5848−5853.
37. Mason, A. B., He, Q.-Y., Adams, T. E., Gumerov, D. R., Kaltashov, I. A., Nguyen, V., and MacGillivray, R. T. A. (2001) Expression, purification, and characterization of recombinant nonglycosylated human serum transferrin containing a C-terminal hexahistidine tag, *Protein Expression Purif. 23*, 142−150.
38. Lebrón, J. A., Bennett, M. J., Vaughn, D. E., Chirino, A. J., Snow, P. M., Mintier, G. A., Feder, J. N., and Björkman, P. J. (1998) Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor, *Cell 93*, 111−123.
39. West, A. P., Jr., Giannetti, A. M., Herr, A. B., Bennett, M. J., Nangiana, J. S., Pierce, J. R., Weiner, L. P., Snow, P. M., and Björkman, P. J. (2001) Mutational analysis of the transferrin

Halbrooks et al.

receiver reveals overlapping HFE and transferrin binding sites, *J. Mol. Biol. 313*, 385−397.

40. Fagerstam, L. G., Frostell-Karlsson, A., Karlsson, R., Persson, B., and Ronnberg, I. (1992) Biospecific interaction analysis using surface plasmon resonance detection applied to kinetic, binding site and concentration analysis, *J. Chromatogr. 597*, 397−410.

41. Malmqvist, M. (1993) Biospecific interaction analysis using biosensor technology, *Nature 361*, 186−187.

42. Giannetti, A. M., and Björkman, P. J. (2004) HFE and transferrin directly compete for transferrin receptor in solution and at the cell surface, *J. Biol. Chem. 279*, 25866−25875.

43. Sipe, D. M., and Murphy, R. F. (1991) Binding to cellular receptors results in increased iron release from transferrin at mildly acidic pH, *J. Biol. Chem. 266*, 8002−8007.

44. Lehrer, S. S. (1969) Fluorescence and absorption studies of the binding of copper and iron to transferrin, *J. Biol. Chem. 244*, 3613−3617.

45. Pakdaman, R., and El Hage Chahine, J. M. (1996) A mechanism for iron uptake by transferrin, *Eur. J. Biochem. 236*, 922−931.

46. Pakdaman, R., Petitjean, M., and El Hage Chahine, J. M. (1998) transferrins—a mechanism for iron uptake by lactoferrin, *Eur. J. Biochem. 254*, 144−153.

47. Zuccola, H. J. (1993) The crystal structure of monoferric human serum transferrin, Georgia Institute of Technology, Atlanta, GA.

48. Hall, D. R., Hadden, J. M., Leonard, G. A., Bailey, S., Neu, M., Winn, M., and Lindley, P. F. (2002) The crystal and molecular structures of diferric porcine and rabbit serum transferrins at resolutions of 2.15 and 2.60 Å, respectively, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 58*, 70−80.

49. Gumerov, D. R., Mason, A. B., and Kaltashov, I. A. (2003) Interlobe communication in human serum transferrin: metal binding and conformational dynamics investigated by electrospray ionization mass spectrometry, *Biochemistry 42*, 5421−5428.

50. Muralidhara, B. K., and Hirose, M. (2000) Anion-mediated iron release from transferrins—the kinetic and mechanistic model for N-lobe of ovotransferrin, *J. Biol. Chem. 275*, 12463−12469.

51. Mizutani, K., Muralidhara, B. K., Yamashita, H., Tabata, S., Mikami, B., and Hirose, M. (2001) Anion-mediated Fe^{3+} release mechanism in ovotransferrin C-lobe—a structurally identified SO_4^{2-} binding site and its implications for the kinetic pathway, *J. Biol. Chem. 276*, 35940−35946.

52. Kurokawa, H., Mikami, B., and Hirose, M. (1995) Crystal structure of diferric hen ovotransferrin at 2.4 Å resolution, *J. Mol. Biol. 254*, 196−207.

53. Yang, A. H. W., MacGillivray, R. T. A., Chen, J., Luo, Y., Wang, Y., Brayer, G. D., Mason, A. B., Woodworth, R. C., and Murphy,

M. E. P. (2000) Crystal structures of two mutants (K206Q, H207E) of the N-lobe of human transferrin with increased affinity for iron, *Protein Sci. 9*, 49−52.

54. Lambert, L. A., Perri, H., and Meehan, T. J. (2005) Evolution of the duplications in the transferrin family of proteins, *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 140*, 11−25.

55. Anderson, B. F., Baker, H. M., Norris, G. E., Rice, D. W., and Baker, E. N. (1989) Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 Å resolution, *J. Mol. Biol. 209*, 711−734.

56. Haridas, M., Anderson, B. F., and Baker, E. N. (1995) Structure of human diferric lactoferrin refined at 2.2 Å resolution, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 51*, 629−646.

57. Peterson, N. A., Anderson, B. F., Jameson, G. B., Tweedie, J. W., and Baker, E. N. (2000) Crystal structure and iron-binding properties of the R210K mutant of the N-lobe of human lactoferrin: implications for iron release from transferrins, *Biochemistry 39*, 6625−6633.

58. Peterson, N. A., Arcus, V. L., Anderson, B. F., Tweedie, J. W., Jameson, G. B., and Baker, E. N. (2002) "Dilysine trigger" in transferrins probed by mutagenesis of lactoferrin: crystal structures of the R210G, R210E, and R210L mutants of human lactoferrin, *Biochemistry 41*, 14167−14175.

59. Young, S. P., Bomford, A., and Williams, R. (1984) The effect of the iron saturation of transferrin on its binding and uptake by rabbit reticulocytes, *Biochem. J. 219*, 505−510.

60. Mason, A. B., He, Q.-Y., Tam, B. M., MacGillivray, R. T. A., and Woodworth, R. C. (1998) Mutagenesis of the aspartic acid ligands in human serum transferrin: lobe-lobe interaction and conformation as revealed by antibody, receptor-binding and iron-release studies, *Biochem. J. 330*, 35−40.

61. Evans, R. W., Crawley, J. B., Garratt, R. C., Grossmann, J. G., Neu, M., Aitken, A., Patel, K. J., Meilak, A., Wong, C., Singh, J., Bomford, A., and Hasnain, S. S. (1994) Characterization and structural analysis of a functional human serum transferrin variant and implications for receptor recognition, *Biochemistry 33*, 12512−12520.

62. Zak, O., Tam, B., MacGillivray, R. T. A., and Aisen, P. (1997) A kinetically active site in the C-lobe of human transferrin, *Biochemistry 36*, 11036−11043.

63. Penhallow, R. C., Brown-Mason, A., and Woodworth, R. C. (1986) Comparative studies of the binding and growth-supportive ability of mammalian transferrins in human cells, *J. Cell. Physiol. 128*, 251−260.

BI0518693

APPENDIX D

Effect of Glycosylation on the Function of a Soluble, Recombinant Form of the Transferrin Receptor[†]

Shaina L. Byrne,[‡] Rachael Leverence,[§] Joshua S. Klein,[ǁ] Anthony M. Giannetti,[ǁ,⊥] Valerie C. Smith,[#]
Ross T. A. MacGillivray,[#] Igor A. Kaltashov,[§] and Anne B. Mason*,[‡]

*Department of Biochemistry, University of Vermont College of Medicine, Burlington, Vermont 05405-0068,
Department of Chemistry, University of Massachusetts at Amherst, Amherst, Massachusetts 01003, Graduate Option in
Biochemistry and Molecular Biophysics, California Institute of Technology, Pasadena, California 91125, and
Department of Biochemistry and Molecular Biology and Centre for Blood Research, University of British Columbia,
Vancouver, British Columbia V6T 1Z3, Canada*

ABSTRACT: Production of the soluble portion of the transferrin receptor (sTFR) by baby hamster kidney (BHK) cells is described, and the effect of glycosylation on the biological function of sTFR is evaluated for the first time. The sTFR (residues 121−760) has three N-linked glycosylation sites (Asn251, Asn317, and Asn727). Although fully glycosylated sTFR is secreted into the tissue culture medium (∼40 mg/L), no nonglycosylated sTFR could be produced, suggesting that carbohydrate is critical to the folding, stability, and/or secretion of the receptor. Mutants in which glycosylation at positions 251 and 727 (N251D and N727D) is eliminated are well expressed, whereas production of the N317D mutant is poor. Analysis by electrospray ionization mass spectrometry confirms dimerization of the sTFR and the absence of the carbohydrate at the single site in each mutant. The effect of glycosylation on binding to diferric human transferrin (Fe$_2$ hTF), an authentic monoferric hTF with iron in the C-lobe (designated Fe$_C$ hTF), and a mutant (designated Mut-Fe$_C$ hTF that features a 30-fold slower iron release rate) was determined by surface plasmon resonance; a small (∼20%) but consistent difference is noted for the binding of Fe$_C$ hTF and the Mut-Fe$_C$ hTF to the sTFR N317D mutant. The rate of iron release from Fe$_C$ hTF and Mut-Fe$_C$ hTF in complex with the sTFR and the sTFR mutants at pH 5.6 reveals that only the N317D mutant has a significant effect. The carbohydrate at position 317 lies close to a region of the TFR previously shown to interact with hTF.

Transferrin (TF) is a bilobal metal binding protein that transports iron to cells. The N- and C-lobes of human TF (hTF)[1] are homologous globular domains that can each bind one atom of ferric iron (Fe^{3+}) in a cleft formed by two subdomains. The iron is coordinated by two tyrosines, a histidine, an aspartic acid residue, and two oxygen atoms from the synergistic carbonate anion, which is anchored by a conserved arginine residue (*1*). Since ferric iron is insoluble in aqueous solution at physiological pH, binding to hTF is an absolute requirement for delivery of iron to cells. Diferric hTF (Fe$_2$ hTF) in the circulation preferentially binds to the extracellular portion of the transferrin receptor (TFR) on the cell surface at neutral pH (∼7.4). Monoferric hTF binds to the TFR an order of magnitude less tightly than Fe$_2$ hTF,

and apo-hTF does not effectively compete with diferric or the two monoferric hTF species for binding to the TFR (*2, 3*). The Fe$_2$ hTF/TFR complex is endocytosed into a clathrin-coated pit forming an endocytic vesicle. The coat disassembles as the endocytic vesicles fuse within the cell. The pH within the endosome is lowered to pH ∼5.6 by a mechanism involving a proton pump which leads to iron release (*4, 5*). Although all of the details related to the release

* Correspondence should be addressed to this author. Phone: (802) 656-0343. Fax: (802) 656-8220. E-mail: anne.mason@uvm.edu.
‡ University of Vermont College of Medicine.
§ University of Massachusetts at Amherst.
ǁ California Institute of Technology.
⊥ Current address: Roche Palo Alto, LLC, Palo Alto, CA 94043.
University of British Columbia.

[1] Abbreviations: WT, wild type; hTF, human serum transferrin that is glycosylated; hTF-NG, human serum transferrin that is nonglycosylated; Fe$_2$ hTF, diferric human serum transferrin; N-His hTF-NG, recombinant nonglycosylated human serum transferrin with an N-terminal hexahistidine tag and a factor Xa cleavage site attached to the amino terminus of the protein; Fe$_C$ hTF, N-His Y95F/Y188F hTF-NG monoferric hTF with iron in the C-lobe; Mut-Fe$_C$ hTF, slowly releasing C-lobe mutant (N-His Y95F/Y188F/R632N/D634N hTF-NG); TFR, transferrin receptor 1; sTFR, recombinant soluble portion of transferrin receptor 1 with an N-terminal hexahistidine tag, a factor Xa cleavage site, and residues 121−760 of the TFR; N251D, N317D, and N727D, sTFR mutants containing Asn → Asp mutations at the indicated positions; DMEM-F12, Dulbecco's modified Eagle's medium−Ham F-12 nutrient mixture; BHK cells, baby hamster kidney cells; CHO cells, Chinese hamster ovary cells; UG, Ultroser G; FBS, fetal bovine serum; BSA, bovine serum albumin; HRP, horseradish peroxidase; TMB, 3,3′,5,5′-tetramethylbenzidine; EDTA, ethylenediaminetetracetic acid; MES, morpholinoethanesulfonic acid; Ni-NTA, nickel nitrilotriacetic acid; DMT-1, divalent metal transporter; ER, endoplasmic reticulum; SPR, surface plasmon resonance; NaN$_3$, sodium azide; ESI MS, electrospray ionization mass spectrometry.

of iron are not clear, it appears that ferric iron may be reduced by a newly described ferrireductase (*6*), and the resulting ferrous iron crosses the endosomal membrane for use by the cell in a process involving the divalent metal transporter, DMT-1 (*7*). Critical to the cycle, apo-hTF remains bound to TFR at acidic pH and is transported back to the plasma membrane surface where it is released to bind more ferric iron. The complete process of iron delivery takes only 2−3 min (*8*).

The ubiquitous TFR (also known as TFR-1) is an inducible 760 amino acid, membrane-bound protein. A constitutively expressed TFR, known as TFR-2, has also been identified (*9*). TFR is initially synthesized as an 86 kDa protein that dimerizes shortly after synthesis forming a homodimer held together by two intermolecular disulfide bonds (Cys89 and Cys98) (*10*, *11*). After exiting the endoplasmic reticulum (ER), N-linked glycosylation occurs at three asparagine residues (Asn251, Asn317, and Asn727) and one O-linked threonine residue (Thr104), yielding a protein with a mass of ∼190 kDa (*10*, *12*). The predicted TFR primary amino acid sequence (*13*, *14*) led to the identification of three distinct regions: a globular extracellular portion which binds hTF (residues 90−760), a hydrophobic membrane-spanning segment (residues 62−89), and the remaining 61 residues which lie within the cytoplasm and contain signaling motifs (*15*). Residues 89−126 of the TFR comprise a stalk separating the extracellular domain from the transmembrane domain.

The crystal structure of a recombinant form of the soluble TFR (sTFR, comprised of residues 121−760) expressed by Chinese hamster ovary cells was determined by Lawrence et al. (*11*). The structure revealed that the extracellular portion of the TFR is comprised of three subdomains: a protease-like domain (resembling amino- and carboxypeptidases), an apical domain, and a helical domain (Figure 1A). Experimentally, the sTFR is easier to work with since it remains soluble in the absence of the detergent required to maintain the full-length TFR in solution. The development of a robust baculovirus/insect cell expression system by Drs. Snow and Bjorkman (including the attachment of a hexa-His tag to the N-terminus for ease of purification) has made recombinant sTFR available to the research community (*16−19*). This expression system also allows production of site-directed mutants to determine the role of specific amino acid residues involved in binding of hTF. Parenthetically, the HFE protein also binds TFR. This protein is defective in individuals suffering from hereditary hemochromatosis in which there is an increase in the intestinal absorption of iron leading to excessive iron stores and iron overload (*20*). Significantly, hTF and HFE compete with each other for binding to TFR, implying that they share recognition sequences on the TFR (*20*, *21*). The availability of the crystal structure of the HFE/sTFR complex identified the amino acid residues involved in binding of HFE to sTFR and, thereby, also provided information with regard to potential binding region(s) of hTF (*22*). An earlier study (*23*) had identified a conserved Arg-Gly-Asp sequence at residues 646−648 in the TFR which is critical to hTF binding and accounts at least in part for the high-affinity interaction. Extensive mutagenesis and binding studies by Giannetti et al. (*17*) identified other specific residues in the TFR crucial to the binding of HFE and of hTF. Importantly, this work showed differential
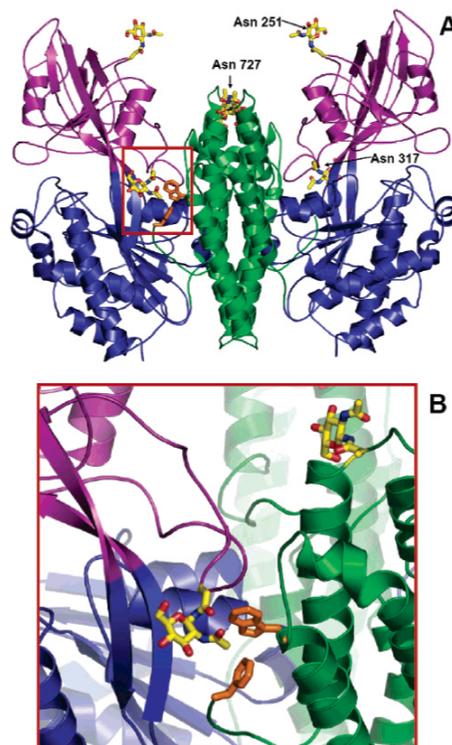
FIGURE 1: Crystal structure of sTFR (PDB code 1CX8) adapted from Lawrence et al. (*11*). (A) Asn-linked glycosylation sites are labeled on one monomer and shown in yellow. The extracellular portion of the TFR is comprised of three subdomains: a protease-like domain resembling amino- and carboxypeptidases (residues 121−188 and 384−606 shown in blue), an apical domain (residues 189−383, purple), and a helical domain (residues 607−760, green). The hydrophobic patch residues Trp641 and Phe760 are in orange. (B) Close-up of the region indicated in the red box in (A) to highlight the residues and the proximity of Asn317 (yellow) to the hydrophobic patch residues Trp641 and Phe760 (orange) involved in hTF binding. The residue in the upper right-hand corner is Asn727.

binding affinities of the sTFR for apo-hTF and Fe_2 hTF as a function of pH. In particular, two TFR residues (Trp641 and Phe760) reside in the helical domain and form a hydrophobic patch (see below).

Earlier work of Aisen and colleagues (initially utilizing TFR isolated from placenta and more recently the recombinant sTFR) was pivotal in establishing the crucial role of the TFR in facilitating iron release from hTF at the appropriate time and place (*24−27*). This group established techniques that unequivocally showed that TFR inhibits iron release from hTF at pH 7.4 and accelerates it at the putative endosomal pH of ∼5.6. More recently, other approaches to map the TFR and hTF interface have been reported (*28*, *29*), complementing the site-directed mutagenesis work from the Bjorkman laboratory mentioned above (*17*). The differential

effect of radiation damage on hTF (or the C-lobe of hTF) and the TFR individually compared to the hTF/TFR complex has been determined, providing a "footprint" that identifies residues in each that are protected by complex formation (*28*). A second approach involved construction of an atomic model obtained by fitting crystal structures of the human N-lobe and rabbit C-lobe into a map of the sTFR (*29*). This cryo-EM study has identified residues in both the sTFR and hTF involved in complex formation. It is proposed that the C-lobe makes contact through the C-I subdomain with the helical region of the sTFR, allowing the C-II subdomain to move freely. It is further suggested that the N-lobe binds TFR through both subdomains in contact with the helical and the protease-like domain on the underside of the TFR, placing the N-lobe between the TFR and the membrane.

Recent experiments provide additional insight into the importance of Trp641 and Phe760 from sTFR in the pH-dependent release of iron from hTF (*30*). The double mutant W641A/F760A sTFR was found to bind Fe_2 hTF with an affinity close to that found for wild-type (WT) sTFR at pH 7.4, 6.3, and 5.6. In contrast, apo-hTF binds this mutant sTFR with a 400−1000-fold lower affinity at pH 6.3 and 5.6, respectively, compared to WT sTFR. Furthermore, the double mutant actually slows iron release from the C-lobe of monoferric hTF by a factor of 2.

As mentioned, TFR has three N-linked glycosylation sites. As shown in Figure 1A, Asn727 is found in the helical domain while Asn 251 and 317 both reside in the apical domain. Of interest is the observation that Asn317 appears to reside close to the hydrophobic patch of the sTFR described above (Figure 1B). Extensive work from the laboratories of Enns, Hunt, and colleagues has established the importance of glycosylation in the proper folding, transport, and insertion of full-length TFR into the plasma membrane (*10*, *31−36*). Protein structure and stability, intracellular trafficking, and localization as well as protection from proteolysis and enhanced solubility are known to be influenced by attachment of carbohydrate (ref *37* and references cited therein). Additionally, different glycoforms have different effects on these properties (*38*).

In the current study we report the development of an expression system using baby hamster kidney (BHK) cells to synthesize and secrete a His-tagged soluble TFR construct similar to that produced previously in insect cells (*18*). The rationale for producing this construct in a different expression system is threefold: (1) to explore the possibility of obtaining higher yields in a system in which we have extensive experience; (2) to provide a different target for crystallization studies since the glycosylation composition is likely to vary in the two expression systems; and (3) to allow us to pursue further mutagenesis studies. We have established a competitive immunoassay to measure the expression levels of the WT and mutant sTFR permitting the optimization of our expression system. Additionally, we have produced mutants in which each of the three asparagine residues have been converted to aspartic acid to prevent glycosylation, and we have analyzed the composition of these constructs by mass spectrometry. We report the binding constants for each of the sTFR glycosylation mutants by using surface plasmon resonance (SPR). We also report the rate constants for the release of iron from the C-lobe of a monoferric hTF (Fe_C hTF) and a mutant of Fe_C hTF (designated Mut-Fe_C hTF)

bound to each sTFR construct. In the case of the Fe_C hTF/sTFR complex we are using a newly developed stopped-flow procedure which provides these rates with greater precision. For the first time, the effect of glycosylation on the biological function of the soluble TFR is revealed.

MATERIALS AND METHODS

Materials. Dulbecco's modified Eagle's medium−Ham F-12 nutrient mixture (DMEM-F12), antibiotic−antimycotic solution (100×), and trypsin were from the GIBCO-BRL Life Technologies Division of Invitrogen. The *Escherichia coli* strain MACHI was also purchased from Invitrogen. Fetal bovine serum (FBS) was obtained from Atlanta Biologicals (Norcross, GA) and was tested prior to use to ensure adequate growth of BHK cells. Ultroser G (UG) is a serum replacement from Pall BioSepra (Cergy, France). The QuikChange mutagenesis kit and pBluescriptII were from Strategene. Ni-NTA resin and the Qiaquick nucleotide removal kit were from Qiagen. The Klenow fragment and buffer were from New England Biolabs. Corning expanded surface roller bottles and Dynatech Immunolon 4 Removawells were obtained from Fisher Scientific. The Hi-Prep 26/60 Sephacryl S-300HR column was from Amersham Pharmacia. Methotrexate from Bedford Laboratories was purchased at a local hospital pharmacy and used for selection of plasmid-containing BHK cells. Centricon 30 microconcentrators, YM-30 ultrafiltration membranes, and spiral cartridge concentrator (CH2PRS) fitted with an S1Y10 cartridge were from Millipore/Amicon. Bovine serum albumin (BSA) was from Sigma. Rabbit anti-mouse immunoglobulin G was from Southern Biological Associates. Immunopure NHS-LC-biotin and immunopure avidin−horseradish peroxidase were from Pierce. The TMB Microwell peroxidase substrate system was obtained from Kirkegaard and Perry Laboratories (Gaithersburg, MD). Human serum TF was purchased from Intergen (Purchase, NY) or from Sigma. The A4A6 monoclonal antibody to TFR was a generous gift from the laboratory of Dr. James Cook at the University of Kansas Medical Center. All other chemicals and reagents were of analytical grade or better.

Preparation of Plasmids. A full-length human TFR cDNA clone was kindly provided by Dr. Caroline Enns (Department of Cell and Developmental Biology, Oregon Health & Science University). The cDNA was engineered for the expression of sTFR that contained the signal peptide of hTF, four amino acids (V-P-D-K) from the N-terminus of hTF, six histidine residues, a factor Xa cleavage site, and the N-terminal region of the TFR beginning at residue 121. A double-stranded synthetic oligonucleotide was formed by hybridizing two overlapping oligonucleotides (Table 1, oligos 1 and 2); aliquots (10 μL of a 2 μg/μL solution) of both oligonucleotides were diluted with 20 μL of Klenow buffer and incubated at 85.0 °C for 30 min, cooled to room temperature, and placed at 4 °C overnight. The overhanging ends were filled in by using the Klenow fragment of *E. coli* DNA polymerase and 10 mM dNTPs. This double-stranded oligonucleotide was purified using the Qiaquick nucleotide removal kit and was used as the forward primer in a PCR reaction together with an internal TFR primer (Table 1, oligo 3) containing an *Mfe*I restriction site. The resulting PCR fragment (coding for a hexa-His-tagged N-terminal sequence and amino acid residues 121−275 of the TFR) was cleaved

Byrne et al.

Table 1: Sequences of the Mutagenic Primers Used in This Study

oligo 1: 5′ -AAA CCC GGG AAG ATG AGG CTC GCC GTG GGA GCC CTG CTG GTC TGC GCC GTC CTA GGG CTG TGT CTG GCT GTC CCT 3′

oligo 2: 5′ TCA GGT CAT CCC AAT ATA AGC GCC TTC CCT CGA TGA ATT CAT GAT GAT GAT GAT GAT GTT TAT CAG GGA CAG CCA GAC A 3′

oligo 3:[a] 5′ CAC ACC AAT TGC ATT TAA 3′

N251D:[b] 5′ TA TAC ACT CCT GTG GAT GGA TCT ATA GTG ATT GTC AGA GC 3′

N317D:[b] 5′ A TTC CCT TCC TTC GAT CAC ACT CAG TTT CCA CCA TCT CGG 3′

N727D:[b] 5′ CAA AAT AAC GGT GCT TTT GAT GAA ACG CTG TTC AGA AAC C 3′

[a] The underlined region represents the *Mfe*I restriction site used in a subsequent cloning step. [b] The mutagenic base is represented by the bold, underlined base.

with *Sma*I and *Mfe*I, made blunt ended with Klenow fragment and 10 mM dNTPs, and cloned into the *Sma*I site of pBluescriptII (pBSSK). Both the full-length TFR construct and the cloned PCR fragment were digested with *Mfe*I and *Xba*I, and the appropriate fragments were isolated by gel electrophoresis and Qiaquick gel extraction purification. The two fragments containing *Mfe*I and *Xba*I overhanging ends were ligated overnight at 14 °C and used to transform *E. coli* strain MACHI. Subsequent DNA sequence analysis revealed a clone with a plasmid containing the correct predicted N-terminal sequence of TFR. The cDNA from the pBSSK clone was cleaved with *Sma*I and *Xba*I; the overhang created by *Xba*I was made blunt using Klenow fragment and 10 mM dNTPs, purified using the Qiaquick gel extraction kit, and ligated into the *Sma*I site of pNUT.

The three N-linked glycosylation sites at positions 251, 317, and 727 in the sTFR in the pNUT vector were mutated from Asn to Asp using the QuikChange mutagenesis procedure (*39, 40*). The mutagenic primers are shown in Table 1. To make the sTFR mutant lacking all three glycosylation sites, the N251D sTFR mutant in the pNUT vector was mutated at position 317 to create a double mutant which was used as the template for the final mutation at position 727. In all cases, the presence of the correct mutation and the absence of unwanted mutations were confirmed by DNA sequence analysis. Transfection of BHK cells and selection with methotrexate were carried out as described in detail previously (*40*).

The expression and purification of sTFR in the baculovirus/insect cell system at the California Institute of Technology have been described in detail previously (*41*). Briefly, the gene encoding residues 121−760 of the TFR was fused 3′ to a gene segment encoding the hydrophobic leader peptide from the baculovirus protein gp67 with a hexa-His tag and a factor Xa cleavage site.

sTFR Production. The protocol for transferring the transfected BHK cells into expanded surface roller bottles has been described previously (*39, 40, 42*). In each production run, the first three batches contain DMEM-F12−10% FBS and were not saved. The subsequent batches (4−7) contained DMEM-F12 with 1% UG + 1 mM butyric acid replacing the FBS and were pooled following addition of 0.02% sodium azide (NaN₃) and stored at 4 °C until purification.

sTFR Assay. The amount of sTFR in the tissue culture medium and at various stages of purification was determined using a specific monoclonal antibody to the TFR in a competitive solid-phase immunoassay as previously described in detail for hTF (*39*). In the assay, a constant amount of biotinylated N-His sTFR was mixed with unlabeled N-His sTFR standards ranging from 16 to 400 ng/well and varying

amounts of sample and added to each well in duplicate. The amount of sTFR in each sample was determined by comparison with the standard curve generated from the unlabeled standards.

sTFR Purification. The purification of N-His sTFR followed a protocol developed for N-His hTF (*40*) in which the pooled batches were concentrated and exchanged into 5 mM Tris, pH 8.0, containing 0.02% NaN₃ using a Millipore spiral wound membrane concentrator fitted with an S1Y10 cartridge. Each sample was filtered through a 0.22 μm Sterivex filter, and Qiagen start buffer (5×) was added to yield a final concentration of 50 mM Tris, pH 7.5, containing 300 mM NaCl, 20 mM imidazole, 10% glycerol, and 0.05% NaN₃. The sample was then loaded onto a Ni-NTA column (1 × 10 cm) at a flow rate of 2 mL/min. The column was attached to a BioCad Sprint system (Applied Biosystems) to allow continuous monitoring of the absorbance at 280 nm, the conductivity, and the pH. After being loaded onto the column and washed with start buffer (1×) until the A_{280} < 0.1, the N-His sTFR was displaced with elution buffer (start buffer containing 250 mM imidazole). Peak fractions were pooled, reduced using YM30 Centricon microconcentrators to less than 2 mL, filtered through a 0.22 μm Millex syringe filter, and loaded onto a Sephacryl S300HR 26/60 column equilibrated and run in 0.1 M NH₄HCO₃ at a flow rate of 1.5 mL/min. Following pooling of appropriate fractions, aliquots were analyzed on a 10% SDS−PAGE gel to verify purity. An identical procedure was used to purify the N-His sTFR obtained from The California Institute of Technology (*15−18*). In this case, the recombinant sTFR was secreted by insect cells into the medium which was shipped to Vermont.

The sTFR contains 11 tryptophans, 24 tyrosines, and 2 cystines, yielding a calculated millimolar extinction coefficient of 96.51 (*43*). The calculated mass of the nonglycosylated sTFR (residues 121−760) is 71726 Da, to which is added 440 Da for the V-P-D-K sequence, 823 Da for the hexa-His tag, and 456 Da for the factor Xa cleavage sequence for a total mass of 73445 Da. From the mass spectrometry results, the monomer mass of the glycosylated TFR is 79760 Da. The concentration (in milligrams per milliliter) for sTFR can be determined from the A_{280} by dividing by 1.21. The production and purification of the recombinant hTF samples, Fe$_C$ hTF and Mut-Fe$_C$ hTF, have been described in detail elsewhere (*40, 44*).

Mass Spectrometry. Electrospray ionization mass spectra of all protein samples were obtained under nondenaturing conditions using a JMS-700 MStation (JEOL, Tokyo, Japan) two-sector mass spectrometer equipped with a standard ESI source. Typically, a 5 μM protein solution in 100 mM

ammonium bicarbonate was continuously injected into the source at a flow rate of 5 μL/min. To avoid in-source oxidation of the protein ions, the spray needle potential was kept below 1.9 kVa. Acceleration voltage was kept at 5 kV, and the nominal resolution was set at 1000. All spectra were recorded by scanning the magnet at a rate of 5 s/decade. Typically, 80−180 scans were averaged for each spectrum to ensure an adequate signal-to-noise ratio. Protein denaturation was carried out by buffer exchanging the original samples into a solution whose pH was adjusted to 2.0 with glacial acetic acid. ESI MS measurements of acid-denatured protein samples were carried out using a QSTAR-XL (PE SCIEX, Framingham, MA) hybrid quadrupole-time-of-flight mass spectrometer equipped with a standard TurboSpray ESI source.

Affinity Measurements Using Surface Plasmon Resonance. A BIACORE 2000 biosensor system (Amersham Biosciences) was used to measure the affinities between the sTFR and hTF samples as described previously (*17, 18*). Binding of injected hTF to sTFR immobilized on the sensor chip results in changes in SPR that are recorded in real time as resonance units (RU) (*45, 46*). An oriented capture method was used to immobilize the purified sTFR samples on a CM5 sensor chip (Amersham Biosciences) by first immobilizing approximately 2200 RU of anti-penta-His antibody (Qiagen) by random amine coupling for all four flow cells. After the surface was blocked with 1 M ethanolamine, pH 8.0, approximately 300 RU of sTFR was immobilized per flow cell, with the exception of flow cell 1 which was used as a blank. In both experiments, the baseline for the N317D sTFR mutant decayed slightly during equilibration of the chip surface in running buffer, possibly accounting for a lower R_{max} relative to other flow cells, but was allowed to stabilize prior to analyte injection. Two separate CM5 chips were required to collect data for all five sTFR samples (BHK WT sTFR, insect cells WT sTFR, BHK N251D, BHK N317D, and BHK N727D). For each analyte (Fe$_2$ hTF, Fe$_C$ hTF, and Mut-Fe$_C$ hTF), a 2-fold dilution series of 10 concentrations preceded and followed by buffer blanks were injected over the flow cells at 70 μL/min at 25 °C in 50 mM Tris buffer, pH 7.4, 150 mM NaCl, and 0.005% P-20 surfactant. The chip surface was regenerated between sample injections with 30 μL of 1 M MgCl$_2$ in running buffer. Primary sensorgram data were preprocessed using the Scrubber software package (Biologic Software Pty.; http://www.biologic.com.au) and globally fitted to 2:1 or 1:1 models in Clamp XP, as previously described (*30, 44, 47*). Sensorgrams corresponding to the highest analyte concentration injections were dropped in the final fitting. The K_D values for the 1:1 models were statistically corrected with a factor of 0.5 such that they may be directly compared to the K_{D1} values for 2:1 models.

Complex Formation and Purification. To prepare Fe$_C$ hTF/sTFR complexes for iron release studies, two different protocols were followed. In the first protocol, a molar excess of Fe$_C$ hTF or Mut-Fe$_C$ hTF was added to 1.0−1.5 mg of WT or mutant sTFR. Following reduction and filtration, the complex was loaded onto a Sephacryl S300HR 26/60 column, equilibrated, and chromatographed as described above. The fractions containing the complex (as confirmed by SDS−PAGE) were pooled and reduced in YM30 Centricon microconcentrators to a nominal concentration of 15 mg/mL. In the second protocol, Fe$_C$ hTF/sTFR complexes

Table 2: Production of Recombinant sTFR and the Glycosylation Mutants of sTFR from BHK Cells

recombinant sTFR	maximum production (mg/L) $\pm$ SD
N-His sTFR WT	34.4 ± 6.1, $n = 4$
N-His sTFR N251D	39.5 ± 8.5, $n = 4$
N-His sTFR N317D	15.5 ± 1.6, $n = 4$
N-His sTFR N727D	30.9 ± 0.9, $n = 2$

were formed by the addition of a slight excess of sTFR and subsequent reduction using a microconcentrator.

Kinetic Rate Studies. The rates of iron release from Mut-Fe$_C$ hTF in complex with sTFR were determined at 25 °C using a QuantaMaster-6 fluorometer from Photon Technology International (PTI), with excitation at 280 nm and emission at 330 nm. A 3 mL cuvette containing 100 mM MES, pH 5.6, 300 mM KCl, and 4 mM EDTA in a volume of 1.8 mL (and a small stir bar to provide mixing) was placed in the fluorometer, and data collection was initiated to establish a baseline. Once equilibrated with respect to temperature, Mut-Fe$_C$ hTF ($\sim$500 nM) or the Mut-Fe$_C$ hTF/sTFR complex (also 500 nM with respect to Mut-Fe$_C$ hTF) was added by using a 25 μL Hamilton syringe through a port directly above the cuvette. The release of iron was monitored at 1 s intervals by measuring the increase in fluorescence. Data for a minimum of four samples were processed and analyzed using Origin 7.5 software and fitted to a single exponential linear equation ($Y = p_1 e^{-x/p_2} + p_3 + p_4 x$), which yielded R^2 values between 0.982 and 0.989.

For experiments with Fe$_C$ hTF, iron release rates were determined using an Applied Photophysics (AP) SX.18MV stopped-flow spectrofluorometer fitted with a 20 μL observation cell with a 2 mm light path and a dead time of 1.1 ms. A monochromator was used for excitation at 280 nm, and the fluorescence emission was measured using a high-pass filter with a 320 nm cutoff. The temperature (25 °C) was kept constant using a circulating water bath. One syringe contained 375 nM (with respect to Fe$_C$ hTF) complex in 1.0 mL of 300 mM KCl (pH $\sim$6.8). The other syringe contained 300 mM KCl, 200 mM MES, pH 5.6, and 8 mM EDTA. Kinetic traces were collected for 50 s intervals a total of six to eight times and averaged. At least three separate samples were averaged for each value reported. Data were analyzed using Origin 7.5 software fit best to a single exponential linear equation (as above); R^2 values varied from 0.996 to 0.999.

RESULTS

sTFR Production. WT sTFR (residues 121−760) and three single point mutants (N251D, N317D, and N727D) were expressed in BHK cells and secreted into the tissue culture medium. Each construct had four amino acids from the N-terminus of hTF and an N-His tag as well as a factor Xa cleavage site. As determined by a competitive immunoassay, the WT and mutant sTFR samples were expressed in this BHK system (Table 2). The results clearly show that production of WT sTFR and the N251D and N727D mutants is comparable, reaching a maximum of 30−40 mg/L, while the N317D mutant was approximately half of this value. Significantly, attempts to express a completely nonglycosylated sTFR construct (the N251D/N317D/N727D triple

Byrne et al.

Table 3: Determination of Masses by Electrospray Mass Spectrometry

recombinant sTFR[a]	post S300 column	monomer (dimer/2) (kDa)	dimer (kDa ± SD)	trimer[b] (kDa ± SD)
N-His sTFR (BHK)	peak	79.7	**159.4 ± 2.2**[c] **159.5 ± 2.8**	NO
	shoulder	80.1	**160.1 ± 2.7**	239.3 ± 2.9
N-His sTFR N251D	peak	77.1	**154.0 ± 1.4** **154.3 ± 1.8**	233.5 ± 3.9[d]
	shoulder	77.1	154.2 ± 1.6 154.2 ± 1.2	233.9 ± 2.5 233.6 ± 2.7
N-His sTFR N317D	peak[e]	77.5	**154.4 ± 2.6** 155.5 ± 4.0	234.1 ± 2.3[d] 235.3 ± 3.9
	shoulder	77.7	155.4 ± 2.8[d]	**235.0 ± 3.3**
N-His sTFR N727D	peak	78.1	**156.2 ± 2.9** **156.0 ± 2.6**	236.4 ± 3.8[d] 235.2 ± 2.8
	shoulder	NM[f]		
N-His sTFR (insect cells)	peak	83.6	**167.2 ± 3.0**	NO

[a] The calculated mass of the sTFR (residues 121−760) is 71726 Da, to which is added 440 Da for the V-P-D-K sequence, 823 Da for the hexa-His tag, and 456 Da for the factor Xa cleavage sequence for a total mass of 73445 Da. The difference is attributed to the carbohydrate (see Results and Discussion). [b] See text. We have determined that the trimer is made up of a dimer of TFR and a molecule of hTF. [c] The bold font indicates the major species present in each sample. [d] The precision of the mass determination on the indicated samples is ±100. All other samples have a precision of ±20. [e] In this sample, a species with a mass of 63470 Da was observed. [f] NM = not measured. NO = not observed.

mutant) were unsuccessful. Both Western blot analysis and our competitive immunoassay confirmed the absence of any secreted sTFR. These results suggest either that glycosylation and secretion are intimately connected or that the secreted product is insoluble.

Following reduction and buffer exchange, the BHK cell medium containing recombinant sTFR was loaded onto a Ni-NTA column to capture the His-tagged constructs. After elution from the nickel column with 250 mM imidazole, final purification involved chromatography using an S300HR 26/60 gel filtration column. In each case, the main protein peak was preceded by a shoulder. Analysis of selected fractions on a 10% SDS−PAGE gel indicated that this shoulder contained a species with a higher molecular mass, as would be expected from the elution profile (see below). On the basis of the assay of the starting material, the yield of the recombinant sTFR constructs was ~60%; this yield is similar to our previously reported recoveries for recombinant hTF-NG (*40*).

Mass Spectrometry. The results of ESI MS analyses of various sTFR samples are summarized in Table 3. Since these experiments were carried out under near-native conditions, formation of multiple adducts resulted in significant broadening of the protein ion peaks in ESI mass spectra (see Figure 2 and Supporting Information) and resulted in rather modest mass measurement precision (ranging from 1.2 to 4.0 kDa for the major species, dimer). Nevertheless, the achieved precision is sufficient to confirm the predicted absence of glycosylation at a single site for each mutant when compared to the WT sTFR with glycans attached to all three sites. For each of the analyzed sTFR samples, the main peak from the S300HR column was always composed of a sTFR dimer,
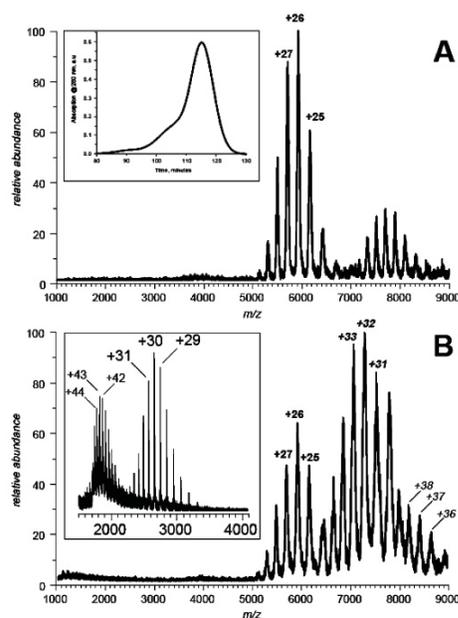

FIGURE 2: Electrospray ionization mass spectra of sTFR N251D size exclusion chromatography peaks: main peak (A) and shoulder (B). Both spectra were acquired under near-native conditions in solution (50 mM ammonium acetate). The elution profile is shown on the inset in panel A. Acid denaturation of the shoulder sample results in emergence of two distinct charge state distributions (inset in panel B), which correspond to the sTFR mutant and hTF (see text for more detail).

with a small amount of a tetrameric species also observed in all four sTFR preparations (Figure 2A). The presence of the tetrameric species in the mass spectra is not surprising, since such low-abundance oligomers are often observed in ESI MS under native conditions and are usually attributed to oligomerization stimulated by increased protein concentration in solution in the ESI interface (*48*). Although sTFR dimers were also observed in the shoulders of the N251D and N317D mutant peaks, a major species in each case had a significantly higher mass than the dimer but lower than the tetramers (Figure 2B). Such species were also observed in the WT sTFR shoulder peak and were only slightly less abundant than the sTFR dimer. Although the measured masses of these species are reasonably close to that of a putative sTFR homotrimer, the limited resolution and accuracy of measurements in the high *m/z* range make it impossible to assign the species solely on the basis of mass. For example, a putative heterotrimer composed of a sTFR dimer and a single TF molecule would have a mass within 2 kDa of the homotrimer. To establish the composition of this species unequivocally, the protein complex in question was denatured with acid, and its monomeric constituents were identified on the basis of their mass measurements in the low *m/z* region (Figure 2B inset). Since protein ions generated under denaturing conditions in solution do not form adducts as readily as those produced under native conditions,

Table 4: SPR Results for Binding of Fe$_2$ hTF, Fe$_C$ hTF, and Mut-Fe$_C$ hTF to WT sTFR and Three Glycosylation Mutants at pH 7.4[a]

transferrin	expt[b]	receptor	$K_{D1} \pm$ SE (nM)	$K_{D2} \pm$ SE (nM)
Fe$_2$ hTF	1	WT sTFR BHK	0.47 ± 0.003	15 ± 0.1
	1	sTFR N251D	0.51 ± 0.003	14 ± 0.1
	1	sTFR N317D	0.59 ± 0.01	19 ± 0.2
Fe$_2$ hTF	2	WT sTFR insect cells[c]	0.65 ± 0.01	10 ± 0.2
	2	sTFR N317D	0.57 ± 0.004	15 ± 0.1
	2	sTFRN727D	0.44 ± 0.003	13 ± 0.1
Fe$_C$ hTF	1	WT sTFR BHK	22 ± 0.3	
	1	sTFR N251D	21 ± 0.3	
	1	sTFR N317D	31 ± 0.5	
Fe$_C$ hTF	2	WT sTFR insect cells	28 ± 0.6	
	2	sTFR N317D	31 ± 0.5	
	2	sTFRN727D	21 ± 0.3	
Mut-Fe$_C$ hTF	1	WT sTFR BHK	27 ± 0.3	
	1	sTFR N251D	31 ± 0.3	
	1	sTFR N317D	41 ± 0.6	
Mut-Fe$_C$ hTF	2	WT sTFR insect cells	27 ± 0.6	
	2	sTFR N317D	31 ± 0.5	
	2	sTFRN727D	23 ± 0.3	

[a] Note that Fe$_2$ hTF binding is described by two dissociation constants and monoferric hTF binding is described by a single K_D. [b] The designations 1 and 2 refer to two different SPR chips needed to run all of the samples. [c] The WT sTFR in experiment 2 was expressed with a baculovirus/insect cell system.
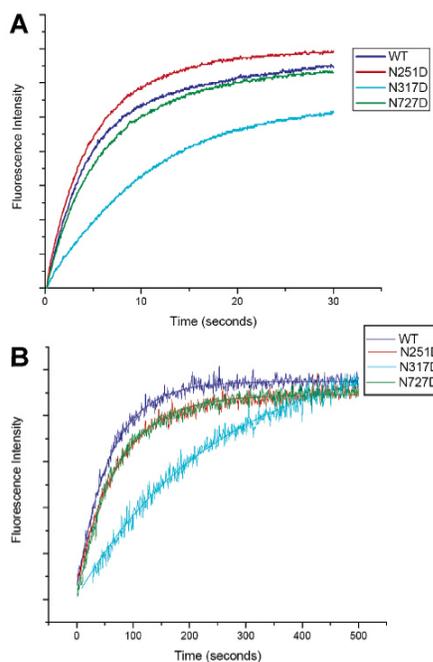
FIGURE 3: Progress curves for the rate of iron release from hTF species in a complex with WT sTFR and the three single glycosylation mutants of sTFR: (A) Fe$_C$ hTF and (B) Mut-Fe$_C$ hTF. The reduced noise observed in the fluorescent profiles of panel A is attributed to the greater sensitivity of the stopped-flow instrument and the averaging of six injections to produce the trace. Note that the time scales are very different in (A) and (B).

significantly higher precision of mass measurement can be easily afforded. The mass spectrum of the acid-denatured protein sample reveals the presence of two distinct ionic species with masses of 79.6 ± 0.1 kDa (charge states +25 through +34) and 78.0 ± 0.1 kDa (charge states +37 through +45). Despite having similar masses, one of the protein species carries a significantly higher number of charges in the gas phase. This provides a rather clear indication that this protein remains more compact in denaturing solution than the other (*49*). This is an expected consequence of the presence of the large number of disulfide bonds in TF (19 total compared to only 2 in the sTFR), which prevent full unfolding of the TF polypeptide chain by imposing multiple conformational constraints. Taken together, measurements of the mass and the extent of multiple charging of the acid-denatured components of protein complexes giving rise to shoulder peaks on size-exclusion chromatograms clearly indicate that they are composed of both sTFR and TF monomers.

Binding of hTF to WT and Mutant sTFR Molecules. To compare the binding characteristics of each mutant sTFR to WT sTFR, equilibrium constants were calculated using an SPR assay to measure binding kinetics in real time (*17, 30*). For these studies, the binding of three different analytes was measured: (1) glycosylated Fe$_2$ hTF, (2) Fe$_C$ hTF, and (3) Mut-Fe$_C$ hTF (selected for its slower rate of iron release; see below) (*44*). We observed that BHK-derived WT sTFR bound to glycosylated Fe$_2$ hTF with the same affinity as WT sTFR produced in the baculovirus/insect system (Table 4). Additionally, and as described previously, the Fe$_2$ hTF binding data were fit best with a bivalent model yielding two K_Ds, whereas Fe$_C$ hTF and Mut-Fe$_C$ hTF were consistent with a 1:1 model yielding a single K_D (*30, 44*). The results indicate that the N251D and N727D sTFR mutants bind hTF

with nearly the same affinities as WT sTFR. In the case of the two Fe$_C$ hTF constructs, the N317D sTFR mutant consistently bound with a lower affinity ($\sim 21.8 \pm 12.0\%$, $n = 4$). In contrast, the K_{D1} for Fe$_2$ hTF binding to the N317D mutant was within experimental error in both experiments. Two experiments were required to analyze all the samples, and the small differences between the two analytes were consistent with the experimental variability ascribed to differences in the individual chips.

Kinetic Studies. Iron release rates were determined by monitoring the increase in fluorescence as iron was released from Fe$_C$ hTF in the presence and absence of sTFR. The fluorescent signal is ascribed to one or more of the five tryptophan residues in the C-lobe of hTF which become(s) unquenched and solvent-exposed as iron is released and the lobe opens (*44, 50*). In previous studies, the fast rate of iron release from Fe$_C$ hTF bound to sTFR yielded large standard deviations in the rate constants that were measured (*19, 30, 44*). In the current work, two strategies were employed to measure the rates with greater precision. In one approach, a stopped-flow instrument from Applied Photophysics was used to acquire the data for the Fe$_C$ hTF/sTFR complexes at pH 5.6. Kinetic curves for each complex are shown in Figure 3A.

Table 5: Kinetics of Iron Release from hTF Mutants with and without sTFR[a]

proteins[b]	k_{obs} ($s^{-1} \times 10^3$) $\pm$ SD	instrument[c]
Fe$_C$ hTF control	3.4 $\pm$ 0.1	Cary
Fe$_C$ hTF + N-His sTFR WT insect cells	262 $\pm$ 44	PTI
Fe$_C$ hTF + N-His sTFR WT BHK	249 $\pm$ 68	PTI
Fe$_C$ hTF + N-His sTFR WT	238 $\pm$ 22	AP
Fe$_C$ hTF + N-His sTFR N251D	204 $\pm$ 9	AP
Fe$_C$ hTF + N-His sTFR N317D	90 $\pm$ 10	AP
Fe$_C$ hTF + N-His sTFR N727D	175 $\pm$ 6	AP
Mut-Fe$_C$ hTF[d]	0.087 $\pm$ 0.017	Cary
Mut-Fe$_C$ hTF + N-His sTFR WT[d]	18 $\pm$ 1	PTI
Mut-Fe$_C$ hTF + N-His sTFR N251D	16 $\pm$ 0.1	PTI
Mut-Fe$_C$ hTF + N-His sTFR N317D	8 $\pm$ 1	PTI
Mut-Fe$_C$ hTF + N-His sTFR N727D	17 $\pm$ 0.1	PTI

[a] Iron release from hTF mutants was determined at 25 °C and pH 5.6 (100 mM MES, 4 mM EDTA, 300 mM KCl). [b] The control is N-His Y95F/Y188F hTF-NG. All constructs are in this background. [c] Measurements were carried out as described in Materials and Methods on a Varian Cary 100 dual beam spectrophotometer, a Photon Technology International QuantaMaster (PTI) spectrofluorometer, or an Applied Photophysics (AP) SX.18MV stopped-flow spectrofluorometer as indicated above. [d] The mutant is N-His Y95F/Y188F/R632N/D634N hTF-NG, and the release rate for the mutant alone has been previously reported (44). In addition, the rate for the mutant in the presence of the sTFR from the insect cells was reported as 25.5 $\pm$ 0.9 $s^{-1} \times 10^{-3}$, $n = 3$ (44).

A second approach to simplify comparisons involved the use of a mutant (designated Mut-Fe$_C$ hTF) with a slower rate of release (30-fold slower in the absence of TFR and 9-fold slower in the presence of TFR) when compared to Fe$_C$ hTF (44). This mutant was originally designed to mimic the composition of a triad of residues found in the C-lobe of lactoferrin which is well-known to have slower iron release rates than ovotransferrin or hTF. Typical release curves from the Mut-Fe$_C$ hTF/sTFR complexes are shown in Figure 3B. A summary of the kinetic rate constants for each sTFR bound to Fe$_C$ hTF and Mut-Fe$_C$ hTF is presented in Table 5. The results clearly show that the N317D sTFR mutant has a 2−3-fold slower release rate than the WT sTFR and the N251D or N727D sTFR mutants.

Since our earlier studies (30, 44) utilized sTFR from baculovirus/insect cells, we wanted to verify the assumption that the two recombinant forms of the WT sTFR are functionally equivalent. Because the amino acid sequence of each WT sTFR differs only by the presence of four extra amino acids (with a mass of 440 Da) preceding the His tag in the BHK-derived sTFR, most of the difference in mass resides in the composition of the carbohydrate. As shown in Table 3, analysis by mass spectrometry reveals that the baculovirus/insect cell derived sTFR is slightly larger (see Discussion). Nevertheless, the SPR data indicate that the binding of Fe$_2$ hTF to the two recombinant sTFR samples is identical (Table 4). Additionally, iron release rates for Fe$_C$ hTF in a complex with either the BHK or insect cell derived sTFR are the same within experimental error (Table 5).

DISCUSSION

By transfecting BHK cells with a mutant in which the three asparagine linkage sites were converted to aspartic acid, we hoped to express a recombinant form of sTFR that lacked glycosylation. We anticipated that production of a nongly-cosylated sTFR would provide a homogeneous preparation for use in mass spectroscopy studies and might aid in crystallization trials. Unfortunately, no nonglycosylated sTFR was secreted into the tissue culture medium of the BHK cells as indicated by an immunoassay and further confirmed by Western blot analysis using a probe for the His tag. This finding differs from expression of hTF, in which glycosylation plays no role in either expression or function (51), but is consistent with a report of the failure of the nongly-cosylated mutant to reach the cell surface in TFR-deficient CHO cells (10). The inability to express the completely nonglycosylated sTFR led us to produce the single sTFR glycosylation mutants individually to facilitate an assessment of the role of each glycosylation site in the expression and function of the sTFR. In our laboratory, the secreted His-tagged sTFR from the baculovirus/insect cell medium obtained from the expression facility at The California Institute of Technology was purified with a final yield of 13.6 $\pm$ 4.4 mg/L, $n = 8$. In the present work, we find that the yield of sTFR from the BHK expression system is comparable, 11.6 $\pm$ 4.3 mg/L, $n = 4$. In each case, the yield of functional sTFR is considerably higher than amounts reported either from the CHO cell system ($\sim$2 mg/L) (11) or from placental preparations (2−6 mg of full-length TFR per placenta) (52, 53).

As previously noted (11, 54, 55), even in the absence of the disulfide linkages in the stalk region, the WT and each of the mutant sTFR constructs form dimers in solution as clearly shown by their behavior during gel filtration chromatography and by mass spectrometry analysis (Table 3). Although it is clear that glycosylation of two of the three sites allows expression, it is unclear whether it might be possible to express a sTFR with a single glycan. Our results make it tempting to speculate that only the carbohydrate at position 317 may be crucial to the production of functional sTFR.

Determining the exact composition of the carbohydrate at each site is extremely challenging because although BHK cells and insect cells attach carbohydrate at the consensus sequences, the composition of the attached carbohydrate is usually variable in both a cell type- and species-dependent manner (ref 51 and references cited therein). Interestingly, in naturally occurring TFR, the complexity of the carbohydrate appears to be specific to the position of the Asn residues in the sequence (32). Thus, it has been reported that human TFR isolated from placenta and TFR expressed in mouse NIH-3T3 cells show similar patterns (32); the Asn251 site featured a complex triantennary, trisialylated carbohydrate with a fucose core (3009 Da), the Asn317 site had a sialylated hybrid oligosaccharide (1874 Da), and Asn727 had a high mannose type oligosaccharide (1866 Da). Our measurements do not allow such a detailed determination of the carbohydrate composition at each site.

Identification of a "trimer" by mass spectrometry analysis is attributed to the presence of a TFR dimer with a single molecule of hTF bound. At the resolution of the analysis the difference in mass between a TFR monomer and a molecule of hTF is indistinguishable. Using acid denaturation and analysis in the low m/z region, we were able to make the distinction unequivocally. As previously reported (39), hTF is present in the serum substitute Ultroser G at a concentration of $\sim$2−4 mg/L. Due to the high affinity of

the hTF/TFR interaction, it is extremely likely that any hTF in the tissue culture medium that acquires iron would bind to the recombinant TFR. Likewise, the complex would be expected to elute from the gel filtration column as a higher molecular weight "shoulder".

Elimination of the carbohydrate at position 251 has little or no effect on the expression, dimerization, complex formation, and release of iron from TF compared to WT. Previous work showed that glycosylation at position 251 is necessary for protection against proteolysis (*34*), although no proteolysis was observed in another study (*36*). The carbohydrate at position 251 was not involved in ligand binding and/or dimerization of the TFR in vivo (*34*). Likewise, we observed no interference with complex formation or dimerization of the N251D mutant, and obviously, protection from proteolysis is not relevant in expression of the secreted soluble portion of the TFR.

In our studies, the N317D sTFR mutant expressed poorly and was more difficult to purify. We note that this mutant appears to be less soluble and/or is possibly more prone to aggregation. In the case of the two Fe_C hTF samples, the absence of carbohydrate at this position causes a small difference of ~20% in the binding affinity measured at pH 7.4 although, interestingly, no significant difference was found for binding of Fe_2 hTF to this mutant (Table 4). Examination of the crystal structure of sTFR (Figure 1) reveals that residue 317 from one monomer is within 4 Å of W641 and F760 on the other monomer. As described in the introduction, these two residues comprise a hydrophobic patch on the TFR that is involved in the binding of hTF. This patch appears to be responsible for stabilization of apo-hTF at acidic pH (*30*). The absence of carbohydrate at position 317 has a small impact on the affinity for Fe_C hTF at pH 7.4. Additionally, the rate of iron release at pH 5.6 is 2–3-fold slower. These results imply that the carbohydrate at this position helps to attain and/or stabilize the conformation of the sTFR in a pH-dependent manner. Consistent with this idea is the observation that the affinity of Fe_2 hTF for the N317D sTFR is the same as that found for WT sTFR and the other mutants (Table 4). Likewise, the change in affinity for the hydrophobic patch double mutant was only observed at pH 6.3 and 5.6. Thus, the finding of equal affinity for Fe_2 hTF at neutral pH is consistent, and the N317D glycosylation may have a role in stabilizing the bound apo-hTF at low pH.

Previous studies indicated that the carbohydrate at position 727 is important in proper folding of the TFR and crucial to transport of the TFR to the plasma membrane (*33, 36*). Obviously, production of the soluble TFR does not require the intracellular trafficking function. In addition, the current work does not suggest that the sTFR is improperly folded since the mutant is expressed at a concentration that is equivalent to the control (Table 2), binds equally well to the various hTF constructs (Table 4), and yields a similar acceleration in the rate of iron release from the C-lobe of hTF (Table 5). Therefore, and within the context of the soluble TFR, the absence of carbohydrate at position 727 has no impact on any of the measured criteria.

In summary, we present data that demonstrate the importance of glycosylation in the expression of the sTFR; no sTFR is expressed when the three sites are mutated to prevent glycosylation. We show that, in contrast to expression of

full-length TFR, only the carbohydrate at position 317 has a significant effect on the expression and iron release rates of the soluble form of the TFR. Additionally, we have shown that the stopped-flow spectrofluorometer is able to capture iron release rates with greater precision and sensitivity than previously used methods.

ACKNOWLEDGMENT

SUPPORTING INFORMATION AVAILABLE

One figure displaying zoomed regions of ESI mass spectra showing ionic signals of sTFR species. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. MacGillivray, R. T. A., Moore, S. A., Chen, J., Anderson, B. F., Baker, H., Luo, Y. G., Bewley, M., Smith, C. A., Murphy, M. E., Wang, Y., Mason, A. B., Woodworth, R. C., Brayer, G. D., and Baker, E. N. (1998) Two high-resolution crystal structures of the recombinant N-lobe of human transferrin reveal a structural change implicated in iron release, *Biochemistry 37*, 7919–7928.

2. Young, S. P., Bomford, A., and Williams, R. (1984) The effect of the iron saturation of transferrin on its binding and uptake by rabbit reticulocytes, *Biochem. J. 219*, 505–510.

3. Mason, A. B., Halbrooks, P. J., James, N. G., Connolly, S. A., Larouche, J. R., Smith, V. C., MacGillivray, R. T. A., and Chasteen, N. D. (2005) Mutational analysis of C-lobe ligands of human serum transferrin: insights into the mechanism of iron release, *Biochemistry 44*, 8013–8021.

4. Aisen, P. (2004) Transferrin receptor 1, *Int. J. Biochem. Cell Biol. 36*, 2137–2143.

5. Rybak, S. L., and Murphy, R. F. (1998) Primary cell cultures from murine kidney and heart differ in endosomal pH, *J. Cell. Physiol. 176*, 216–222.

6. Ohgami, R. S., Campagna, D. R., Greer, E. L., Antiochos, B., McDonald, A., Chen, J., Sharp, J. J., Fujiwara, Y., Barker, J. E., and Fleming, M. D. (2005) Identification of a ferrireductase required for efficient transferrin-dependent iron uptake in erythroid cells, *Nat. Genet. 37*, 1264–1269.

7. Gunshin, H., and Hediger, M. A. (2002) The divalent metal-ion transporter (DCT1/DMT1/Nramp2), in *Molecular and Cellular Iron Transport* (Templeton, D. M., Ed.) pp 155–173, Marcel Dekker, New York.

8. Teeters, C. L., Lodish, H. F., Ciechanover, A., and Wallace, B. A. (1986) Transferrin and apotransferrin: pH-dependent conformational changes associated with receptor-mediated uptake, *Ann. N.Y. Acad. Sci. 463*, 403–407.

9. Kawabata, H., Yang, R., Hirama, T., Vuong, P. T., Kawano, S., Gombart, A. F., and Koeffler H. P. (1999) Molecular cloning of transferrin receptor 2. A new member of the transferrin receptor-like family, *J. Biol. Chem. 274*, 20826–20832.

10. Yang, B., Hoe, M. H., Black, P., and Hunt, R. C. (1993) Role of oligosaccharides in the processing and function of human transferrin receptors. Effect of the loss of the three *N*-glycosyl oligosaccharides individually or together, *J. Biol. Chem. 268*, 7435–7441.

11. Lawrence, C. M., Ray, S., Babyonyshev, M., Galluser, R., Borhani, D. W., and Harrison, S. C. (1999) Crystal structure of the ectodomain of human transferrin receptor, *Science 286*, 779–782.

12. Enns, C. A., Clinton, E. M., Reckhow, C. L., Root, B. J., Do, S.-I., and Cook, C. (1991) Acquisition of the functional properties of the transferrin receptor during its biosynthesis, *J. Biol. Chem. 266*, 13272−13277.

13. McClelland, A., Kuhn, L. C., and Ruddle, F. H. (1984) The human transferrin receptor gene: genomic organization, and the complete primary structure of the receptor deduced from a cDNA sequence, *Cell 39*, 267−274.

14. Schneider, C., Owen, M. J., Banville, D., and Williams, J. G. (1984) Primary structure of human transferrin receptor deduced from the mRNA sequence, *Nature 311*, 675−678.

15. Kuhn, L. C., McClelland, A., and Ruddle, F. H. (1984) Gene transfer, expression, and molecular cloning of the human transferrin receptor gene, *Cell 37*, 95−103.

16. Lebrón, J. A., and Bjorkman, P. J. (1999) The transferrin receptor binding site on HFE, the class I MHC-related protein mutated in hereditary hemochromatosis, *J. Mol. Biol. 289*, 1109−1118.

17. Giannetti, A. M., Snow, P. M., Zak, O., and Bjorkman, P. J. (2003) Mechanism for multiple ligand recognition by the human transferrin receptor, *PLoS Biol. 1*, 341−350.

18. West, A. P., Jr., Giannetti, A. M., Herr, A. B., Bennett, M. J., Nangiana, J. S., Pierce, J. R., Weiner, L. P., Snow, P. M., and Bjorkman, P. J. (2001) Mutational analysis of the transferrin receptor reveals overlapping HFE and transferrin binding sites, *J. Mol. Biol. 313*, 385−397.

19. Zak, O., and Aisen, P. (2003) Iron release from transferrin, its C-lobe, and their complexes with transferrin receptor: Presence of N-lobe accelerates release from C-lobe at endosomal pH, *Biochemistry 42*, 12330−12334.

20. Enns, C. A. (2001) Pumping iron: the strange partnership of the hemochromatosis protein, a class I MHC homolog, with the transferrin receptor, *Traffic 2*, 167−174.

21. Giannetti, A. M., and Bjorkman, P. J. (2004) HFE and transferrin directly compete for transferrin receptor in solution and at the cell surface, *J. Biol. Chem. 279*, 25866−25875.

22. Bennett, M. J., Lebrón, J. A., and Bjorkman, P. J. (2000) Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor, *Nature 403*, 46−53.

23. Dubljevic, V., Sali, A., and Goding, J. W. (1999) A Conserved RGD (Arg-Gly-Asp) motif in the transferrin receptor is required for binding to transferrin, *Biochem. J. 341*, 11−14.

24. Bali, P. K., Zak, O., and Aisen, P. (1991) A new role for the transferrin receptor in the release of iron from transferrin, *Biochemistry 30*, 324−328.

25. Bali, P. K., and Aisen, P. (1991) Receptor-modulated iron release from transferrin: Differential effects on N- and C-terminal sites, *Biochemistry 30*, 9947−9952.

26. Bali, P. K., and Aisen, P. (1992) Receptor-induced switch in site−site cooperativity during iron release by transferrin, *Biochemistry 31*, 3963−3967.

27. Aisen, P. (1992) Entry of iron into cells: a new role for the transferrin receptor in modulating iron release from transferrin, *Ann. Neurol. 32* (Suppl.), S62−S68.

28. Liu, R. T., Guan, J. Q., Zak, O., Aisen, P., and Chance, M. R. (2003) Structural reorganization of the transferrin C-lobe and transferrin receptor upon complex formation: The C-lobe binds to the receptor helical domain, *Biochemistry 42*, 12447−12454.

29. Cheng, Y., Zak, O., Aisen, P., Harrison, S. C., and Walz, T. (2004) Structure of the human transferrin receptor-transferrin complex, *Cell 116*, 565−576.

30. Giannetti, A. M., Halbrooks, P. J., Mason, A. B., Vogt, T. M., Enns, C. A., and Bjorkman, P. J. (2005) The molecular mechanism for receptor-stimulated iron release from the plasma iron transport protein transferrin, *Structure 13*, 1613−1623.

31. Reckhow, C. L., and Enns, C. A. (1988) Characterization of the transferrin receptor in tunicamycin-treated A431 cells, *J. Biol. Chem. 263*, 7297−7301.

32. Hayes, G. R., Williams, A. M., Lucas, J. J., and Enns, C. A. (1997) Structure of human transferrin receptor oligosaccharides: Conservation of site-specific processing, *Biochemistry 36*, 5276−5284.

33. Williams, A. M., and Enns, C. A. (1991) A mutated transferrin receptor lacking asparagine-linked glycosylation sites shows reduced functionality and an association with binding immunoglobulin protein, *J. Biol. Chem. 266*, 17648−17654.

34. Hoe, M. H., and Hunt, R. C. (1992) Loss of one asparagine-linked oligosaccharide from human transferrin receptors results in specific cleavage and association with the endoplasmic reticulum, *J. Biol. Chem. 267*, 4916−4923.

35. Hunt, R. C., Riegler, R., and Davis, A. A. (1989) Changes in glycosylation alter the affinity of the human transferrin receptor for its ligand, *J. Biol. Chem. 264*, 9643−9648.

36. Williams, A. M., and Enns, C. A. (1993) A region of the C-terminal portion of the human transferrin receptor contains an asparagine-linked glycosylation site critical for receptor structure and function, *J. Biol. Chem. 268*, 12780−12786.

37. Wang, C., Eufemi, M., Turano, C., and Giartosio, A. (1996) Influence of the carbohydrate moiety on the stability of glycoproteins, *Biochemistry 35*, 7299−7307.

38. Rudd, P. M., Joao, H. C., Coghill, E., Fiten, P., Saunders, M. R., Opdenakker, G., and Dwek, R. A. (1994) Glycoforms modify the dynamic stability and functional activity of an enzyme, *Biochemistry 33*, 17−22.

39. Mason, A. B., He, Q.-Y., Adams, T. E., Gumerov, D. R., Kaltashov, I. A., Nguyen, V., and MacGillivray, R. T. A. (2001) Expression, purification, and characterization of recombinant nonglycosylated human serum transferrin containing a C-terminal hexahistidine tag, *Protein Expression Purif. 23*, 142−150.

40. Mason, A. B., Halbrooks, P. J., Larouche, J. R., Briggs, S. K., Moffett, M. L., Ramsey, J. E., Connolly, S. A., Smith, V. C., and MacGillivray, R. T. A. (2004) Expression, purification, and characterization of authentic monoferric and apo-human serum transferrins, *Protein Expression Purif. 36*, 318−326.

41. Lebrón, J. A., Bennett, M. J., Vaughn, D. E., Chirino, A. J., Snow, P. M., Mintier, G. A., Feder, J. N., and Bjorkman, P. J. (1998) Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor, *Cell 93*, 111−123.

42. Mason, A. B., Funk, W. D., MacGillivray, R. T. A., and Woodworth, R. C. (1991) Efficient production and isolation of recombinant amino-terminal half-molecule of human serum transferrin from baby hamster kidney cells, *Protein Expression Purif. 2*, 214−220.

43. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci. 4*, 2411−2423.

44. Halbrooks, P. J., Giannetti, A. M., Klein, J. S., Bjorkman, P. J., Larouche, J. R., Smith, V. C., MacGillivray, R. T. A., Everse, S. J., and Mason, A. B. (2005) Composition of pH sensitive triad in C-lobe of human serum transferrin. Comparison to sequences of ovotransferrin and lactoferrin provides insight into functional differences in iron release, *Biochemistry 44*, 15451−15460.

45. Fagerstam, L. G., Frostell-Karlsson, A., Karlsson, R., Persson, B., and Ronnberg, I. (1992) Biospecific interaction analysis using surface plasmon resonance detection applied to kinetic, binding site and concentration analysis, *J. Chromatogr. 597*, 397−410.

46. Malmqvist, M. (1993) Biospecific interaction analysis using biosensor technology, *Nature 361*, 186−187.

47. Myszka, D. G., and Morton, T. A. (1998) CLAMP: a biosensor kinetic data analysis program, *Trends Biochem. Sci. 23*, 149−150.

48. Peschke, M., Verkerk, U. H., and Kebarle, P. (2004) Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method, *J. Am. Soc. Mass Spectrom. 15*, 1424−1434.

49. Kaltashov, I. A., and Eyles, S. J. (2002) Studies of biomolecular conformations and conformational dynamic by mass spectrometry, *Mass Spectrom. Rev. 21*, 37−71.

50. Lehrer, S. S. (1969) Fluorescence and absorption studies of the binding of copper and iron to transferrin, *J. Biol. Chem. 244*, 3613−3617.

51. Mason, A. B., Miller, M. K., Funk, W. D., Banfield, D. K., Savage, K. J., Oliver, R. W. A., Green, B. N., MacGillivray, R. T. A., and Woodworth, R. C. (1993) Expression of glycosylated and nonglycosylated human transferrin in mammalian cells. Characterization of the recombinant proteins with comparison to three commercially available transferrins, *Biochemistry 32*, 5472−5479.

52. Turkewitz, A. P., Amatruda, J. F., Borhani, D., Harrison, S. C., and Schwartz, A. L. (1988) A high yield purification of the human transferrin receptor and properties of its major extracellular fragment, *J. Biol. Chem. 263*, 8318−8325.

53. Hemadi, M., Kahn, P. H., Miquel, G., and Hage Chahine, J. M. (2004) Transferrin's mechanism of interaction with receptor 1, *Biochemistry 43*, 1736−1745.

54. Borhani, D. W., and Harrison, S. C. (1991) Crystallization and X-ray diffraction studies of a soluble form of the human transferrin receptor, *J. Mol. Biol. 218*, 685−689.

55. Fuchs, H., Lücken, U., Tauber, R., Engel, A., and Gessner, R. (1998) Structural model of phospholipid-reconstituted human transferrin receptor derived by electron microscopy, *Structure 6*, 1235−1243.

APPENDIX E

Published as Win MN, Klein JS, Smolke CD (2006) Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay. *Nucleic Acids Res.* **34**: 5670-82. My contribution to this work was in assisting with the development of a novel method for analyzing the binding properties of RNA aptamers to small-molecules by surface plasmon resonance.

Codeine-binding RNA aptamers and rapid determination of their binding constants using a direct coupling surface plasmon resonance assay

Maung Nyan Win, Joshua S. Klein and Christina D. Smolke*

Division of Chemistry and Chemical Engineering 1200 E. California Boulevard, MC 210-41 California Institute of Technology Pasadena, CA 91125, USA

ABSTRACT

RNA aptamers that bind the opium alkaloid codeine were generated using an iterative *in vitro* selection process. The binding properties of these aptamers, including equilibrium and kinetic rate constants, were determined through a rapid, high-throughput approach using surface plasmon resonance (SPR) analysis to measure real-time binding. The approach involves direct coupling of the target small molecule onto a sensor chip without utilization of a carrier protein. Two highest binding aptamer sequences, FC5 and FC45 with K_d values of 2.50 and 4.00 μM, respectively, were extensively studied. Corresponding mini-aptamers for FC5 and FC45 were subsequently identified through the described direct coupling Biacore assays. These assays were also employed to confirm the proposed secondary structures of the mini-aptamers. Both aptamers exhibit high specificity to codeine over morphine, which differs from codeine by a methyl group. Finally, the direct coupling method was demonstrated to eliminate potential non-specific interactions that may be associated with indirect coupling methods in which protein linkers are commonly employed. Therefore, in addition to presenting the first RNA aptamers to a subclass of benzylisoquinoline alkaloid molecules, this work highlights a method for characterizing small molecule aptamers that is more robust, precise, rapid and high-throughput than other commonly employed techniques.

INTRODUCTION

Codeine is a naturally-occurring opium alkaloid, part of the larger class of benzylisoquinoline alkaloids (BIAs), found in the opium poppy, *Papaver somniferum*, and constitutes ∼0.5% of opium (1). It is one of the most widely used narcotic drugs for the treatment of mild to moderate pain, diarrhea and cough with relatively low side effects (2). Despite its extensive medical applications, codeine is often abused for its euphoric and depressant effects as well as to prevent opiate withdrawal (3). Due to increasing misuse, codeine has been incorporated into workplace and military drug testing programs, and a screening and confirmation cut-off concentration of 40 μg/l has been suggested for federally-mandated testing in oral fluid by the Substance Abuse and Mental Health Services Administration (3). Therefore, a sensor system that can precisely measure the concentration of codeine and effectively discriminate against its structural analogues is highly desired.

Aptamers are nucleic acid molecules that bind ligands with high specificity and affinity (4). There is increasing interest in utilizing aptamers as the target recognition elements in various sensing applications (5–8). In addition to the drug detection applications of a codeine-binding aptamer, there are other potential biotechnology applications for this aptamer. Codeine is a member of the BIA family and is a key product metabolite in the opium alkaloid biosynthesis pathway (9). The BIAs comprise a structurally diverse group of pharmacologically important compounds (10) and efforts are ongoing to engineer microbial and plant hosts for the production of some of the important BIA intermediates in the codeine synthesis pathway, such as (S)-reticuline and thebaine (9–11).

Aptamers to BIA molecules may prove to be useful tools for such engineering efforts. Recent research has highlighted the application of aptamers as components of synthetic and naturally-occurring cellular sensors and switches (12–17), which can regulate enzyme levels in response to small molecule ligand concentrations. Therefore, aptamer-based cellular sensors may be generated to act as 'intelligent' regulatory tools for metabolic engineering efforts to provide dynamic regulation of gene expression at specific enzymatic steps so that pathway fluxes are rewired to enable the accumulation of desired intermediate metabolites, which has proven to be difficult to achieve in natural plant

*To whom correspondence should be addressed. Tel: +1 626 395 2460; Fax: +1 626 568 8743; Email: smolke@cheme.caltech.edu

hosts (9). A codeine-binding aptamer may be used to construct tools, such as synthetic riboswitches that can be employed to redirect flux through an engineered BIA metabolic pathway or in setting up rapid functional screens of pathway variants. In addition, while aptamers have been developed to several of the far upstream metabolites in this pathway, such as dopamine (18) and tyrosine (19), they have not yet been developed against any BIA compounds, which harbor bulky, nitrogen-containing ring structures. Prior work has demonstrated that aptamers to specific molecules within a family of compounds may be used to design doped libraries for the selection of aptamers to similar compounds within that family from smaller library sequence spaces (19), and thus codeine aptamers would be potentially useful for selecting aptamers to diverse BIA molecules.

This work describes the generation of novel RNA aptamers to the small molecule codeine and highlights a robust, high-throughput assay method for measuring small molecule-aptamer binding properties. RNA aptamers that bind codeine with high affinities were selected from a combinatorial library containing a 30 nt randomized region using an iterative *in vitro* selection procedure or SELEX (Systematic Evolution of Ligands by EXponential enrichment) (20,21). The binding properties of the generated codeine aptamers were measured by surface plasmon resonance (SPR) through a real-time binding assay (Biacore), similar to previously reported methods (22,23) where the small molecule ligand is directly coupled to a sensor chip through chemical modification of the ligand, eliminating the need to use protein linkers between the target small molecule and the sensor surface as described in other methods (24–26). This direct coupling method limits potential non-specific interactions or binding artifacts arising from the presence of the linker protein observed in previous studies (24,25), which may alter the determined binding affinities. Therefore, this method may provide a more accurate assessment of small molecule-aptamer binding affinities since the measured interaction more closely mimics the binding environment of the *in vitro* selection process.

MATERIALS AND METHODS

DNA template library preparation

A random DNA library was generated through PCR using the following oligonucleotide sequences: a 59 nt DNA template 5′-GGGACAGGGCTAGC(N$_{30}$)GAGGCAAAGCTTCCG-3′, primer1 5′-TTC*TAATACGACTCACTATA*GGGACAGGGC-TAGC-3′ and primer2 5′-CGG<u>AAGCTT</u>TGCCTC-3′. All DNA synthesis was performed by Integrated DNA Technologies, Inc. The template contains a 30 nt randomized region flanked by two fixed primer-binding regions (Figure 1A). Primer1 contains a 17 nt T7 promoter sequence (italic). NheI and HindIII restriction endonuclease sites (underlined) were included in primer1 and primer2, respectively, for cloning of aptamer sequences.

Codeine coupling and affinity chromatography matrix preparation

Approximately 300 mg of epoxy-activated Sepharose 6B (GE Healthcare) was hydrated and incubated with 2.5 mM codeine in coupling buffer [0.05 M Na$_2$PO$_4$ (pH 13)] overnight at 37°C according to the manufacturer's instructions. The coupled medium was washed three times with 2 ml of coupling buffer to remove uncoupled codeine. The medium was then incubated overnight with 1 M Tris–HCl (pH 8) at 40°C to block any remaining active groups. Finally, the medium was washed with a solution containing 0.1 M NaOAc (pH 4) and 0.5 M NaCl followed by a second solution containing 0.1 M Tris–HCl (pH 8) and 0.5 M NaCl. The wash was repeated twice and the matrix was resuspended in 10 mM Tris–HCl (pH 8) and stored at 4°C. The codeine affinity chromatography matrix was prepared by packing the coupled medium (500 μl) into a column following the manufacturer's instructions (Pierce). The packed column was washed with 10 column volumes of binding buffer [250 mM NaCl, 20 mM Tris–HCl (pH 7.4) and 5 mM MgCl$_2$] and equilibrated prior to the selection process.

Initial RNA library pool preparation

The initial DNA library pool was generated by PCR conducted for 12 cycles on a mixture (100 μl) containing 20 pmol DNA template, 300 pmol each primer1 and primer2, 200 μM each dNTPs, 1.6 mM MgCl$_2$, and 10 U *Taq* DNA polymerase (Roche). This DNA library pool (~1.2 × 10^{14} molecules) was transcribed into an initial RNA library pool by incubating overnight at 37°C in the presence of 40 mM Tris–HCl (pH 7.9), 16 mM MgCl$_2$, 10 mM DTT, 2 mM spermidine, 3 mM each rNTPs, 50 μCi [α-^{32}P]UTP (GE Healthcare), 500 U RNase inhibitor and 50 U T7 RNA polymerase (New England Biolabs). The DNA template was subsequently degraded by incubating the reaction mixture with 10 U of DNase I (Invitrogen) at 37°C for 15 min. The unincorporated nucleotides were removed with a NucAway spin column (Ambion) following the manufacturer's instructions and binding buffer was added to the flow-through RNA to bring the total volume up to 500 μl.

In vitro selection of codeine-binding aptamers

Prior to incubation with the codeine-modified affinity column, the RNA pool was denatured at 70°C for 3 min and allowed to renature at room temperature for 30 min. To eliminate RNA molecules that non-specifically bind to the column matrix, the initial pool was first incubated with an unmodified column. The flow-through fraction from this incubation was subsequently transferred to a codeine-modified affinity column and incubated for 45 min. Following the incubation period, the affinity column was washed with 10 column volumes of binding buffer for cycles 1 to 5 to remove unbound RNAs. This wash volume was increased 10 column volumes for each of the subsequent cycles. Bound RNA was eluted with 7 column volumes of 5 mM codeine in binding buffer. The eluted RNA was recovered by ethanol precipitation in the presence of 20 μg/ml glycogen. Reverse transcription and cDNA amplification (15 PCR cycles) were performed in a single step using 200 U of SuperScript III reverse transcriptase (Invitrogen) and 5 U of *Taq* DNA polymerase in a 50 μl reaction volume. One-fifth of this DNA library was transcribed into an RNA library pool for the subsequent selection cycle. A total of 15 selection cycles were carried out during the *in vitro* selection process.

A.

TTC*TAATACGACTCACTATA* (GGGACAGG<u>GCTAGC</u>) (N)$_{30}$ (GAGGCA<u>AAGCTT</u>CCG)
        T7 promoter   5' constant region   3' constant region

B.

Clone	Sequence		
FC21 (2)	GGGACAGGGCTAGC	AAAAGGGTGGTTGAAGGGACAGCTGGTGTG	GAGGCAAAGCTTCCG *
A25 (3)	GGGACAGGGCTAGC	ACAAGAATTAGGGTCGGGAAATGGTGTGTG	GAGGCAAAGCTTCCG *
C4 (2)	GGGACAGGGCTAGC	CACAAGTGTGAAGGGATGGGAGTAGTGGTG	GAGGCAAAGCTTCCG *
C9 (2)	GGGACAGGGCTAGC	AAGAATAGGATGTGGGTAAAGGTGCTGGTG	GAGGCAAAGCTTCCG *
C12 (3)	GGGACAGGGCTAGC	ACATGGAGGCTTATAGGGATTCGTGCTGGG	GAGGCAAAGCTTCCG *
FC5 (1)	GGGACAGGGCTAGC	AGTAGGATTGGGTGAGGGGATGTGCTGTG	GAGGCAAAGCTTCCG *
B10 (1)	GGGACAGGGCTAGC	AGTAGGATTAGGGTGAGGGGATGTGCTGTG	GAGGCAAAGCTTCCG
A28 (1)	GGGACAGGGCTAGC	ACATTGTGGGAAAGGGAATTGAGTGTGGTG	GAGGCAAAGCTTCCG
B11 (1)	GGGACAGGGCTAGC	ACATTGAGGGAAAGGGAATTGAGTGTGGTG	GAGGCAAAGCTTCCG *
FC1	GGGACAGGGCTAGC	CACGAAATGGGTGAAGGGAAACGTGGTGGG	GAGGCAAAGCTTCCG
FC3	GGGACAGGGCTAGC	ACCAAAAATAGGGGTAAGGGCATGGGGGTG	GAGGCAAAGCTTCCG *
FC13	GGGACAGGGCTAGC	AGGGTAAGGGGATTGGAGTAGTGCCGTGTG	GAGGCAAAGCTTCCG *
FC17	GGGACAGGGCTAGC	GGACAAGAAGTGGGTAAGGGAATCCGTGGG	GAGGCAAAGCTTCCG *
FC23	GGGACAGGGCTAGC	CAATAAATAAGGCGAAGTAAGGGATGGGGTG	GAGGCAAAGCTTCCG *
FC27	GGGACAGGGCTAGC	TACTAATGTACGCACTAAGGGATTGGGGTG	GAGGCAAAGCTTCCG *
FC33	GGGACAGGGCTAGC	GAAAGCGGTTTGGGAAAGTAAAGGGTGGTG	GAGGCAAAGCTTCCG
FC34	GGGACAGGGCTAGC	TACAGAATAAGCGAATTAAGGGTTGGGGTG	GAGGCAAAGCTTCCG *
FC36	GGGACAGGGCTAGC	AAAGTGAGGGTTATGGGGATACGTGGCGTG	GAGGCAAAGCTTCCG
FC41	GGGACAGGGCTAGC	ATTAGGGTAATCGATCAAGAGGGAGTGGTG	GAGGCAAAGCTTCCG
FC45	GGGACAGGGCTAGC	TTAGTGCTATGTGAGAAAAGGGTGTGGGGG	GAGGCAAAGCTTCCG *
A2	GGGACAGGGCTAGC	ACGTTAGGATGAGGGTAATGGCGTTGTAGAAGA	GAGGCAAAGCTTCCG
A3	GGGACAGGGCTAGC	GTAATAAGTAGGGAAAGGGTTCCCGCTGGG	GAGGCAAAGCTTCCG *
A5	GGGACAGGGCTAGC	TTTAAAGTGAGGGGTTATGGGCAGTGTGGT	GAGGCAAAGCTTCCG
A7	GGGACAGGGCTAGC	TTTTAAGCACAATAACAGGGTGGGGATGGT	GAGGCAAAGCTTCCG
A15	GGGACAGGGCTAGC	ACCATTAGGGATTATCCAACGGGGGGTGTG	GAGGCAAAGCTTCCG
A20	GGGACAGGGCTAGC	CTATAGTGAGGCTATTAAGGGTTGTGGGGG	GAGGCAAAGCTTCCG *
A22	GGGACAGGGCTAGC	AGTTGAATAGGGTTGGAGAAAGGACGTGGT	GAGGCAAAGCTTCCG
A23	GGGACAGGGCTAGC	TTATTTAGGGTTGGAGGGTAGTTAGCGGTG	GAGGCAAAGCTTCCG
A30	GGGACAGGGCTAGC	TAATGAAGGGCAAGGGAATAGTGGCTAGGG	GAGGCAAAGCTTCCG
B1	GGGACAGGGCTAGC	GAGTAAAAAGGGTTGGGAAAATCGCATGGT	GAGGCAAAGCTTCCG
B2	GGGACAGGGCTAGC	GCAGAACAGAGGGTAGGGAATTTGCGTGTG	GAGGCAAAGCTTCCG *
B4	GGGACAGGGCTAGC	TCAGAACGCTAGATTAGGATGTGGGTGGTG	GAGGCAAAGCTTCCG
B6	GGGACAGGGCTAGC	AAAAGGGTGGTTGAAGGGACAGCTGGTGTG	GAGGCAAAGCTTCCG
B8	GGGACAGGGCTAGC	TACAATAGGGCAATTAATGGGGAGTGTGTG	GAGGCAAAGCTTCCG
B9	GGGACAGGGCTAGC	ATCGGTGTAGGGAAGGGATATGATGTGGTG	GAGGCAAAGCTTCCG *
B10	GGGACAGGGCTAGC	AGTAGGATTAGGGTGAGGGGATGTGCTGTG	GAGGCAAAGCTTCCG
B12	GGGACAGGGCTAGC	AGCGGTAAGGGTGGGGAGAATGGTGCTGTG	GAGGCAAAGCTTCCG *
C1	GGGACAGGGCTAGC	ATAGCATGGAGCGACTATGCGTTGATGGGT	GAGGCAAAGCTTCCG
C3	GGGACAGGGCTAGC	CGTTGTAACGGTGAATTTAGGGTAAGGGGG	GAGGCAAAGCTTCCG
C7	GGGACAGGGCTAGC	CCGTCCCTATAGTGAGTCGTATTAGAACGG	AAGCTTCCG
C10	GGGACAGGGCTAGC	TTTACAGTGAAAAATTAAGGGAAGGGGGTG	GAGGCAAAGCTTCCG
C13	GGGACAGGGCTAGC	AACAAATTTAGGGAGGGATTCGTTGTGGTG	GAGGCAAAGCTTCCG
C14	GGGACAGGGCTAGC	GTTAGGGATGGGGATAACCACATGCGTGGG	GAGGCAAAGCTTCCG
C15	GGGACAGGGCTAGC	AAAGGTTAAGGGTAGGGGATACTGCTGGTG	GAGGCAAAGCTTCCG *
C18	GGGACAGGGCTAGC	AGAAGTACAAGGGTAGGGATTTGGTGCGGG	GAGGCAAAGCTTCCG
C23	GGGACAGGGCTAGC	AGCAAATTGGGAAAGGGATACGCTGTGTGG	GAGGCAAAGCTTCCG *
C24	GGGACAGGGCTAGC	GTATCATTCGCTTCTTATCAGGGTAGGGGG	GAGGCAAAGCTTCCG

Figure 1. Aptamer clone sequences. (**A**) DNA template from which the initial RNA pool was generated. (**B**) Sequences of clones from the final aptamer pool. The codeine-binding properties of the sequences marked with an asterisk were characterized by the described direct coupling SPR assay. The number in parenthesis represents the frequency of a particular clone in the sequenced pool.

At the tenth cycle, a counter-selection against morphine was performed by eluting the bound RNA with 3 column volumes of 5 mM morphine in binding buffer prior to elution with codeine. Only RNA eluted with codeine was used to make the input DNA library pool for the subsequent selection cycle. Following the reverse transcription step of cycles 11, 12 and 13, an error-prone PCR was performed in a mutagenic buffer containing 40 pmol each primer1 and primer2, 7 mM $MgCl_2$, 50 mM KCl, 10 mM Tris–HCl (pH 8.3), 0.2 mM dGTP, 0.2 mM dATP, 1 mM dCTP, 1 mM dTTP and 0.5 mM $MnCl_2$. One-fifth of the error-prone PCR product from each of these cycles was used as the input DNA library pool for the subsequent selection cycle.

Aptamer library sequence analysis

The DNA pool from cycle 15 was amplified by PCR and cloned into a plasmid using the NheI and HindIII restriction sites present in the fixed regions of the aptamer sequence and the plasmid construct. This plasmid library was transformed into an electrocompetent *Escherichia coli* strain, DH10B (Invitrogen; F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lacZ*ΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ- *rps*L *nup*G). Subcloning was confirmed by colony PCR, and a total of 58 positive colonies were sequenced by Laragen, Inc. The resulting sequences were aligned using the ClustalX sequence alignment program.

Qualitative binding affinity assay

Radiolabeled RNA was prepared from ∼1 μg of the final DNA pool (cycle 15) in the presence of 40 mM Tris–HCl (pH 7.9), 14 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 3 mM each rAGC mix, 150 μM rUTP, 50 μCi [α-^{32}P]UTP, 40 U RNase inhibitor and 50 U T7 RNA polymerase. After allowing the transcription reaction to proceed for 3 h at 37°C, 5 U of DNase I were added to the mixture and the reaction was incubated for 15 min. The unincorporated nucleotides were removed with a NucAway spin column and the flow-through RNA was divided equally into two volumes. One of the radiolabeled RNA pools was incubated with a codeine-modified column, whereas the other pool was incubated with an unmodified column. After a 15 min incubation, each column was washed with 3 column volumes of binding buffer followed by elution with 7 column volumes of 5 mM codeine in binding buffer. The eluted RNA from each column was separated by electrophoresis on an 8% polyacryamide/7 M urea gel in 1× Tris–borate buffer. The gel was dried and the recovered radiolabeled RNA was imaged on a FX phosphorimager (BioRad).

Quantitative direct coupling small molecule-aptamer binding assay

A CM5 sensor chip was primed with RNase-free water followed by preconditioning with a 50 mM sodium hydroxide, 0.1% hydrochloric acid, 0.1% (w/v) SDS, 0.085% phosphoric acid solution prior to immobilization of codeine onto the chip surface. The chip was subsequently activated with a 0.2 M N-ethyl-N′-(dimethylaminopropyl)carbodiimide (EDC), 0.05 M N-hydroxysuccinimide (NHS) solution. An amine surface was created by injecting a solution of 0.1 M 1,8-diaminooctane dissolved in 50 mM sodium borate (pH 8.5)

over the activated sensor chip at 5 μl/min for 10 min. In order to couple codeine to the amine surface, codeine was modified at its hydroxyl group with a succinimidyl group by placing 10 mM codeine in a pyridine solution containing 40 mM disuccinimidyl carbonate and 40 mM 4-dimethylamino pyridine. This modification reaction was allowed to take place for 30 min and the reaction mixture was subsequently diluted with 100 mM sodium borate (pH 7.0) in a 1:1 v/v ratio. Trenbolone (Figure 2B), a small molecule structurally distinct from codeine, was modified in the same manner for use as a background response. The modified trenbolone and codeine molecules were separately coupled onto flow cells 1 and 2 of the sensor chip, respectively, by alternating injections for 7 min at 5 μl/min for a total of 28 min for each molecule. After ligand coupling, the chip was deactivated with 1 M ethanolamine (pH 8.5) and primed twice with binding buffer.

RNA samples (initial pool, final pool and randomly-selected individual sequences from the final pool) were prepared for Biacore analysis using the Ampliscribe T7 High Yield Transcription Kit (Epicentre) following the manufacturer's instructions. Samples were sequentially injected over the sensor surface for 1.5 min at 5 μl/min with a 2 min dissociation time. For each sample, various RNA concentrations were injected by serially diluting samples from 48 to 0.375 μM along with two blank samples containing just binding buffer for use as double referencing. After each run, the surface was regenerated with 10 mM EGTA for 2 min at 5 μl/min. The raw data were processed and analyzed to determine the binding constant for each aptamer using Scrubber (Biologic Software, Pty, Australia, http://www.cores.utah.edu/interaction/).

Isocratic affinity elution and specificity assays

Radiolabeled FC5 and FC45 RNA were prepared using the Ampliscribe T7 High Yield Transcription Kit with minor modifications to the manufacturer's instructions (3 mM each rATP, rCTP, rUTP, 150 μM rGTP and 50 μCi [α-^{32}P]GTP). After 3 h of incubation, DNase I was added to the transcription mixture and the reaction was incubated at 37°C for 15 min. Unincorporated nucleotides were removed with a NucAway spin column.

Isocratic affinity elution assays were performed on radiolabeled FC5 and FC45 as described previously (27,28). The binding affinities to codeine and morphine in solution were determined using the following equation: $K_d = [L_{el}] \times (V_{el} - V_n)/(V_e - V_{el})$, where L_{el} is the free ligand concentration used to elute bound RNA, V_{el} and V_e are the elution volumes for RNA in the presence and absence of free ligand in binding buffer, respectively, and V_n is the column void volume.

Specificity assays were performed by equally dividing the flow-through radiolabeled FC5 and FC45 into three Sepharose columns (300 μl) modified with codeine. After a 30 min incubation, each column was washed with 7 column volumes of binding buffer. Columns were then eluted with a 5 mM solution of the different targets (codeine, morphine or thebaine) in binding buffer, and 5 column volumes of the elution were collected. Collected samples were added to 10 ml of Safety-Solve scintillation liquid (Reseach Products
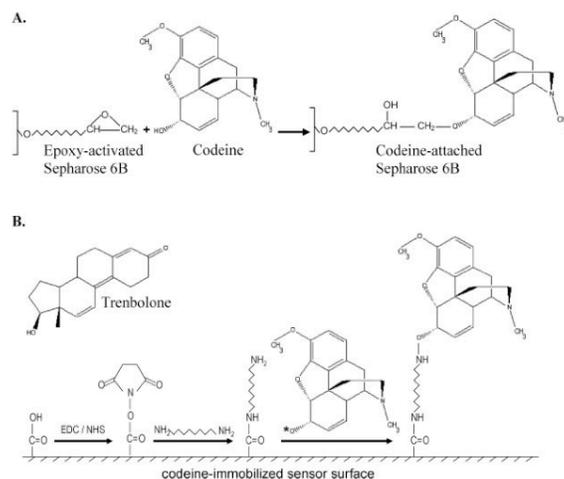
Figure 2. Schematics of the codeine-immobilized surfaces used in the *in vitro* selection process and SPR binding property assay. Illustration of the chemistries used for codeine coupling to the (**A**) Sepharose matrix and (**B**) Biacore CM5 sensor chip surface. Note that the codeine-immobilized sensor surface more closely mimics that of the affinity matrix used during the aptamer selection process versus coupling methods that employ a protein linker. The asterisk next to the oxygen group of codeine in (B) represents a succinimidyl group (the same group that is covalently attached to the carboxyl group of the sensor surface after EDC/NHS activation), which reacts with the amine group of the 1,8-diaminooctane linker. Codeine is thereby immobilized onto the chip surface through the same functional group used to attach it to the affinity matrix during the selection process. Trenbolone, the negative control molecule, is immobilized to the chip surface through the same chemistry and its structure is shown in (B).

International Corp.) and radioactivity levels were measured on a liquid scintillation counter (Beckman Coulter).

Truncation experiments

Two full-length aptamers with the lowest determined K_d values were truncated primarily into four different sequences containing distinct regions of their parent sequences: (i) the random region (Ran), (ii) the cloning region (Cln), (iii) the random region and the $5'$ constant terminus (L) and (iv) the random region and the $3'$ constant terminus (R). Predicted secondary structures formed by these truncated sequences were examined using mfold (29) and RNAstructure (http://rna.chem.rochester.edu/RNAstructure.html). Sequences that adopt well-defined secondary structures were selected for subsequent K_d determination through the described small molecule-aptamer binding affinity SPR assay.

Structural probing assay

Structural probing of the FC5 and FC45 full-length aptamers was performed using a lead ion cleavage assay as described by Berens *et al*. (30) with the following slight modifications. $5'$ end labeled RNA was incubated in binding buffer containing 0–250 μM codeine and 0.5 mM lead (II) acetate. After a 15 min incubation, the cleavage reactions were stopped by adding 0.5 mM EDTA and 1 μg/μl glycogen and the cleaved RNA was recovered by ethanol precipitation. Radiolabeled RNA was also subject to RNase T1 cleavage (Ambion) and alkaline hydrolysis (Ambion) following the manufacturer's instructions to be used as ladders. The recovered RNA

samples were separated by electrophoresis on a 10% polyacryamide/8 M urea gel in 1× Tris–borate buffer. The gel was dried and the RNA cleavage patterns were imaged with an FX phosphorimager (BioRad).

Dopamine aptamer binding assay

For direct coupling of dopamine to the sensor surface, a CM5 sensor chip was activated with EDC/NHS as described above. Following the EDC/NHS activation step, a 10 mM dopamine, 50 mM sodium borate (pH 8.5) solution was injected over the activated sensor surface for 30 min at 5 μl/min to couple dopamine to the surface through its amino group. This is the same chemistry used in the selection of dopamine-binding aptamers described by Mannironi *et al*. (18). After dopamine immobilization, the sensor surface was deactivated with 1 M ethanolamine for 10 min at 5 μl/min to block the remaining unreacted succinimidyl groups. The previously selected dopamine-binding dopa2 RNA aptamer (18) was synthesized using a similar transcription procedure as described above. Various concentrations of this RNA sample were injected over the dopamine-coupled sensor surface and concentration-dependent binding responses were recorded and subsequently analyzed for binding properties as described previously.

For indirect coupling of dopamine to the sensor surface through a BSA protein linker, a CM5 sensor chip was activated with EDC/NHS as described above. Following the EDC/NHS activation step, BSA was injected over the activated surface at 5 μl/min until a signal of 12 500 response units (RU) was reached. A 0.2 M EDC and 0.1 M dopamine

solution was injected over the BSA-immobilized surface for 30 min at 5 μl/min to couple dopamine to BSA. This chemistry couples dopamine to BSA through the same functional group as in the direct coupling chemistry. The remaining steps in the indirect coupling method are identical to those used in the direct coupling method.

RESULTS

Selection of codeine-binding RNA aptamers

A slightly modified *in vitro* selection procedure was used to isolate codeine-binding RNA aptamers from a library of RNA molecules containing a 30 nt random region flanked by constant primer-binding sequences (Figure 1A). Aptamers were selected on a codeine affinity column, which was made by immobilizing codeine to the epoxy-activated agarose through its hydroxyl group (Figure 2A). To enhance the stringency of the selection process, the wash volume was increased incrementally from cycles 6 to 15. To increase the specificity of the selected pool, a counter-selection with a 5 mM morphine solution was performed at cycle 10 prior to elution with codeine. In addition, a total of three error-prone PCR steps were carried out for the DNA template pools of cycles 11, 12 and 13, respectively, to potentially introduce sequences that are of slightly diverse nucleotide composition and search a larger sequence space for higher affinity binders. After cycle 15, the enriched pool was cloned and approximately 60 colonies were sequenced.

Sequence analysis of codeine aptamers reveals fairly conserved motifs

Sequence analysis of the selected clones revealed five sets of completely identical sequences, 2 with 1 nt difference (FC5/B10 and A28/B11), and the rest unique sequences (Figure 1B). Nucleotide deletions or insertions were observed in a few clones. No single consensus sequence that is conserved among the entire population of the selected clones was discovered. However, several fairly conserved but short sequences were found to exist in many clones. For instance, the motif AAGGG is present in over 50% of the sequenced aptamer population. In addition, most clones contain stretches of G's and/or UG's, which suggests that these stretches may be critical to codeine-binding. Similarly, none of the selected clones exhibit obvious predicted structural similarities. It was also observed that most of the sequences do not possess a predominant predicted structure based on analysis with folding programs, such as mfold and RNAstructure, and may adopt different motifs depending on the inclusion or exclusion of part or all of the constant regions in the structural analysis. As a result, further analysis through truncation experiments was conducted to determine the crucial sequences involved in codeine-binding for two aptamers (FC5 and FC45) that exhibited the highest binding affinities of those analyzed.

### Qualitative assessment of codeine-binding affinity of the enriched final pool

The codeine-binding affinity of the final pool was qualitatively assessed by monitoring eluted levels of the radiolabeled aptamer pool using codeine affinity chromatography.

Radiolabeled RNA from the enriched pool was incubated with codeine-modified and unmodified columns. The eluted RNA from each column was run on a polyacrylamide gel and visualized with a phosphorimager (Supplementary Figure 1). Significantly stronger radioactive signals were detected in the sample eluted from the codeine affinity column than that eluted from the unmodified column, indicating that the RNA aptamers in the final pool are highly enriched in codeine-binding affinity.

Determination of small molecule-aptamer binding constants using a direct coupling SPR assay

Quantitative assessment of the codeine-binding properties of the final pool, the initial pool and several aptamers from the final pool was performed using a modified SPR assay developed on a Biacore 2000. Previous studies where SPR was used to determine binding affinities between aptamers and non-protein targets involved the use of BSA or biotin/streptavidin as intermediate linkers between the sensor surface and the target molecules (24,26). Here we employ a direct coupling approach, similar to a previously described method (22,23), in which the small molecule target is directly coupled to the sensor surface without a supporting intermediate, such as BSA or biotin/strepavidin. Previous direct coupling strategies have used target molecules that contain an amine group (22,23), which is a commonly used functional group in Biacore sensor chip immobilization strategies. However, since codeine does not contain an amine group, a chemical modification strategy was developed to directly couple codeine to the chip surface through its hydroxyl group. In this coupling strategy codeine is first modified at its hydroxyl group with an amine-reactive succinimidyl group. This chemical modification enables codeine molecules to readily react with the amine groups attached to the activated chip surface (Figure 2B). Trenbolone was also immobilized onto the sensor surface in the same manner and used as a negative control molecule. Following the immobilization of codeine and trenbolone in their respective flow cells of the sensor chip, serial dilutions of RNA samples were injected into these flow cells. The response detected from the trenbolone-immobilized flow cell was used as the background subtraction in evaluating the binding constants. An equilibrium binding curve was generated from concentration-dependent binding response data for each sample to determine the corresponding K_d value.

The binding data from the SPR assay supports the qualitative binding data obtained from the chromatography-based assay. The data indicate that there was little to no detectable binding (Figure 3A) between the initial pool and codeine, whereas the final pool bound codeine with significant binding responses (Figure 3B). The overall K_d value of the final pool was evaluated to be ∼15 μM, whereas that of the initial pool was estimated to be in the high millimolar range. This latter value is only an estimate as no binding curve could be established for the initial pool due to its insufficient binding response. Therefore, codeine-binding affinity of the final pool was enhanced over 1000-fold from that of the initial pool. The K_d values of the analyzed aptamer clones are listed in Table 1. Several of the aptamer sequences have K_d values that are much lower than that of the enriched final pool.
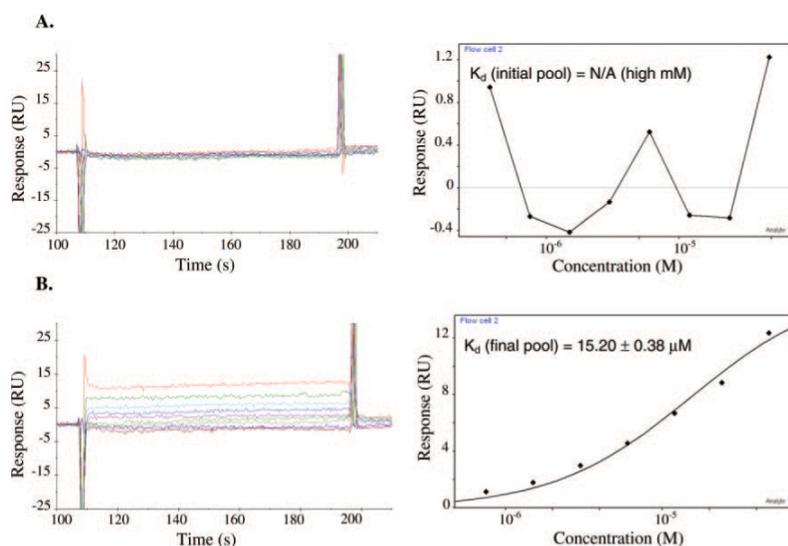
Figure 3. Concentration-dependent codeine-binding responses (left) and the corresponding equilibrium binding curve (right) of (**A**) the initial pool and (**B**) the enriched final pool. Codeine was coupled to the sensor chip as described. Serial dilutions of the appropriate RNA sample were injected across the sensor surface and binding responses were recorded over time. Kinetic rate constants were determined by examining the rate of change of binding response when the RNA samples were initially injected over the surface until equilibrium responses were reached (k_{on}) and when a solution lacking the RNA sample was injected over the surface once equilibrium levels were bound to the chip surface (k_{off}). Equilibrium binding constants (K_d) were determined by plotting the equilibrium binding response versus the RNA sample concentration and calculating the corresponding RNA concentration at which half of the maximal response was achieved. Binding responses were adjusted for background binding by subtracting responses of the corresponding RNA samples determined from a trenbolone-coupled sensor surface.

Table 1. Codeine-binding affinities of the full-length aptamer sequences as determined from the direct coupling SPR assay

RNA sample	K_d (μM)	RNA sample	K_d (μM)	RNA sample	K_d (μM)
Final pool	15.20 ± 0.38	FC27	10.90 ± 0.95	B11	5.80 ± 0.29
Initial pool	N/A (high mM)	FC34	28.00 ± 1.42	B12	8.80 ± 0.44
FC3	28.60 ± 0.96	FC45	2.50 ± 0.06	C4	78.00 ± 4.15
FC5	4.00 ± 0.13	A3	11.50 ± 0.27	C9	9.17 ± 0.33
FC13	14.50 ± 0.58	A20	13.00 ± 0.65	C12	8.88 ± 0.39
FC17	43.70 ± 1.23	A25	7.23 ± 0.34	C15	7.67 ± 0.26
FC21	23.60 ± 1.22	B2	4.75 ± 0.32	C23	8.18 ± 0.23
FC23	19.10 ± 1.49	B9	5.77 ± 0.36		

Two of the highest binding aptamers FC45 and FC5, with K_d values of 2.50 ± 0.06 μM and 4.00 ± 0.13 μM, respectively, were subject to further characterization studies (Figure 4). Despite their similar affinities for codeine, FC5 and FC45 may form different binding pockets since their corresponding mini-aptamers adopt different predicted secondary structures supported by structural studies described in a later section. In addition, FC5 and FC45 exhibit fairly different binding kinetics (Table 2), where the latter has faster kinetics (both binding and dissociation) than that of the former. Some clones, such as FC3, FC13, FC34 and C9 have observed dissociation constants on the same order as that of FC5, while other clones, such as FC23, A3, A20, B11, C15 and C23 exhibit similar dissociation kinetics to FC45 (data not

shown). The kinetic data of the modified FC5 and FC45 sequences discussed in later sections are also reported in Table 2.

The affinities of the two highest binding aptamers, FC5 and FC45, to codeine in solution were also determined through a standard isocratic affinity elution method (27,28). This control enables the comparison of the surface-based binding affinities determined with the described SPR assays to the solution-based affinities. The determined solution binding affinity of FC45 ($K_d = 4.5$ μM) was very similar to its surface binding affinity ($K_d = 2.5$ μM), whereas FC5 was determined to bind free codeine with an ~10-fold lower affinity ($K_d = 47$ μM) than that to surface-immobilized codeine ($K_d = 4.0$ μM). For a given aptamer-ligand pair, the binding
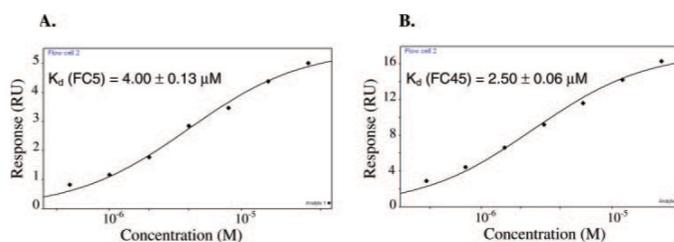
153

Figure 4. Equilibrium codeine-binding curves of (**A**) FC5 and (**B**) FC45.

Table 2. Dissociation rate constants (k_{off}) for codeine binding of the final pool, FC5, FC45 and their corresponding truncated sequences

RNA sample	k_{off} (1/s) %	RNA sample	k_{off} (1/s) %
Final pool	$7.62 \times 10^{-3} \pm 5.81$	Initial pool	N/A
FC5	$6.50 \times 10^{-3} \pm 3.78$	FC45	$1.14 \times 10^{-2} \pm 3.35$
FC5L	$6.70 \times 10^{-3} \pm 2.80$	FC45L	$1.03 \times 10^{-2} \pm 3.61$
FC5L-S1	$6.65 \times 10^{-3} \pm 2.79$	FC45L-S1	$6.84 \times 10^{-3} \pm 3.98$
FC5L-S2	$6.54 \times 10^{-3} \pm 2.44$	FC45L-S2	$2.43 \times 10^{-3} \pm 3.29$
FC5L-S3	$4.79 \times 10^{-3} \pm 5.00$	FC45L-S3	$2.69 \times 10^{-3} \pm 2.23$

The corresponding association rate constant (k_{on}) is equivalent to k_{off}/K_d.

affinities for free target in solution and a target immobilized onto a solid support may differ, as has been observed in previous studies (19,31,32). For the aptamers studied here, FC5 shows differing affinities for free and immobilized codeine, whereas FC45 exhibits similar binding affinities.

Assays reveal distinct specificities of the codeine-binding aptamers to other benzylisoquinoline alkaloid targets

The ability of FC5 and FC45 to distinguish between three similar BIA molecules, codeine, thebaine and morphine, was determined using a chromatography-based assay. Radiolabeled RNA aptamers were eluted with codeine, morphine and thebaine, which are all closely related structural analogues (Figure 5A). Eluted FC5 and FC45 demonstrated ~4- and 6-fold increases in radioactivity counts, respectively (Figure 5B), when eluted with codeine versus morphine. The semi-quantitative molecular specificities of these aptamers were supported by isocratic affinity elution experiments in which the solution affinities of these aptamers were determined and observed to differ by similar magnitudes. The solution affinity for FC45 was determined to be ~4.5 μM to codeine and 25 μM to morphine, whereas the solution affinity for FC5 was determined to be ~47 μM to codeine and 212 μM to morphine. These results demonstrate that the single morphine counter-selection performed during the *in vitro* selection process was effective at enhancing the specificity of the aptamers in the final pool to codeine over morphine. While aptamers that discriminate between molecules that differ by a single methyl group have been described previously for purine alkaloid targets (33,34), these results indicate that aptamers can exhibit this

level of molecular discrimination in spite of the presence of the bulky 4 six-membered rings in the BIA targets examined here.

These assays also demonstrate that these two aptamers exhibit differing specificities to thebaine. The eluted FC5 exhibited nearly a 2-fold increase in radioactivity counts when eluted with thebaine versus codeine, whereas FC45 exhibited an ~30% decrease in signal. These results indicate that FC5 exhibits higher specificity for thebaine over codeine, whereas FC45 exhibits higher specificity for codeine over thebaine. It should be noted that during the selection process codeine was coupled to the Sepharose column in such a way that there was no differentiable functional group between codeine and thebaine. With the attachment chemistry used in these studies through the functional group at C5, these two molecules exhibit conformational differences in that the former has one double bond in the C5-six-membered ring, whereas the latter contains two (Figure 5A). These results suggest that aptamers can potentially perform molecular discrimination at the level of conformation, as the difference between these two targets is at the level of torsional structure of the ring backbone.

Characterization of mini-aptamers that demonstrate binding affinities similar to the full-length aptamers

Truncation experiments were systematically performed on the full-length FC5 and FC45 aptamers to identify minimal aptamer domains, or mini-aptamers. Various truncated aptamer sequences were characterized for their codeine-binding properties. Truncated sequences that form well-defined secondary structures as predicted by mfold or RNAstructure were selected for further analysis. The described SPR small molecule-aptamer binding assays were employed to determine the codeine-binding affinities of these truncated sequences.

An FC5 mini-aptamer was identified by characterizing three truncated sequences of the FC5 full-length aptamer. The codeine-binding properties of the random region (FC5Ran), which is the N30 region of the aptamer library; the cloning region (FC5Cln), which includes the random region, most of the 3′ constant terminus, and part of the 5′ constant terminus; and FC5L, which includes the random region and the 5′ constant terminus, were analyzed using the described SPR binding assay. No binding was observed between FC5Ran and the codeine-immobilized sensor

A.

codeine   thebaine   morphine

B.

Specificity Profile
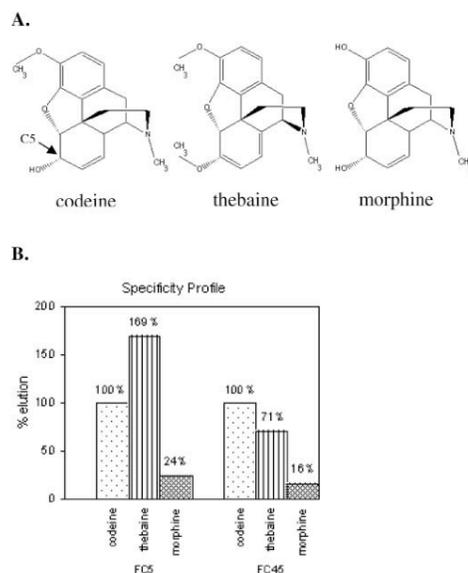
Figure 5. The FC5 and FC45 aptamers exhibit differing specificities to BIA structural analogues. (**A**) Structures of the three BIA molecules, codeine, thebaine and morphine, used in examining aptamer specificity. (**B**) Specificity elution profiles of the FC5 and FC45 aptamers. Radiolabeled aptamers were incubated with a codeine-modified Sepharose matrix. The bound aptamers were subsequently eluted with the different BIA targets and radioactivity levels in the eluted fractions were measured. Radioactivity levels were normalized with respect to values obtained from the codeine elutions for each aptamer.

surface, indicating that the FC5 random region is not sufficient for the codeine-binding properties of this aptamer. FC5Cln demonstrated a significantly reduced affinity to codeine ($K_d = 39.50 \pm 2.27$ µM), suggesting that the remainder of the 5′ constant terminus of FC5 may play an important role in the formation of the correct binding pocket for codeine. FC5L binds codeine with an affinity similar to that of its full-length (59 nt) parent sequence ($K_d = 4.55 \pm 0.14$ µM) despite its significantly reduced length (41 nt). These results indicate that FC5L, referred to as FC5 mini-aptamer, contains the necessary and sufficient sequence within FC5 for binding codeine (Figure 6A and C).

An FC45 mini-aptamer was identified by characterizing two truncated sequences of the FC45 full-length aptamer. The codeine-binding properties of the cloning region (FC45Cln), which includes the random region, most of the 3′ constant terminus, and part of the 5′ constant terminus; and FC45L, which includes the random region and the 5′ constant terminus, were analyzed using the described SPR binding assay. FC45Ran, harboring the N30 region of the library, was not analyzed in this set of truncation experiments, as there was no well-defined secondary structure predicted for this sequence by mfold or RNAstructure.

FC45Cln did not exhibit binding to codeine, suggesting that the codeine-binding pocket was not correctly formed within the secondary structure adopted by this sequence. However, FC45L (44 nt) binds codeine with an affinity (2.59 ± 0.09 µM) that is almost identical to that of the full-length FC45 sequence (Figure 6B and D). Therefore, this FC45 mini-aptamer includes the sequence within FC45 required to form the correct binding pocket for codeine in contrast to that of FC45Cln.

These truncation experiments support the importance of the formation of the correct binding pocket for aptamer molecular recognition capabilities. In addition, both FC5 and FC45 mini-aptamers lack the 3′ constant terminal sequence, indicating that the 3′ terminus is not involved in binding codeine. Secondary structure predictions from mfold and RNAstructure indicate that the 3′ terminus forms a small hairpin (Supplementary Figure 3C), isolating itself from the remaining sequences of FC5 and FC45. The proposed secondary structures of the FC5 and FC45 mini-aptamers (Figure 6) are supported by the structural modification and structural probing experiments described in the next section.

Characterization of modified mini-aptamer sequences supports the proposed secondary structures

The proposed secondary structures of the FC5 and FC45 mini-aptamers do not possess a strong base stem (Figure 6) in comparison to other reported aptamer structures. For instance, the tetracycline minimer (30) has a base stem that is comprised of five base-pairs, which contribute to the stability of the overall secondary structure of the minimer. Sequences lacking strong or stabilized base stems may adopt a number of possible secondary structures, whereas a stabilized base stem can significantly reduce presumed structural variability and therefore restrict a given aptamer sequence to adopt a very few, and in some cases just one, distinct structures. Therefore, the proposed secondary structures of the FC5 and FC45 mini-aptamers may be evaluated by examining the binding properties of these aptamers modified with stabilized base stems.

The base stems of the mini-aptamers were modified with an extension of GC base pairs to stabilize the proposed structures of these mini-aptamers. The FC5 mini-aptamer (FC5L) was stabilized by extending the existing 3 bp stem with two GC base pairs (Figure 7B), based on the assumption that a few nucleotides present on each end of the original mini-aptamer are unessential for codeine-binding. Similarly, the FC45 mini-aptamer (FC45L) was stabilized by extending the base stem formed by the 5′-CUU and 3′-GGG pairing with two GC base pairs (Figure 8B), excluding several nucleotides from the 5′ end. Following the modification, the structures of these stabilized mini-aptamers were further analyzed in mfold using the DotPlot Partition Function, which confirms these structures to be the most favorable ones to adopt among others. The codeine-binding properties of the resulting mini-aptamers, referred to as FC5L-S1 and FC45L-S1, respectively, were determined using the described SPR assay. FC5L-S1 and FC45L-S1 were determined to bind codeine with K_d values of 5.51 ± 0.23 µM and 4.18 ± 0.48 µM, respectively (Supplementary Figure 2A and B). These results indicate
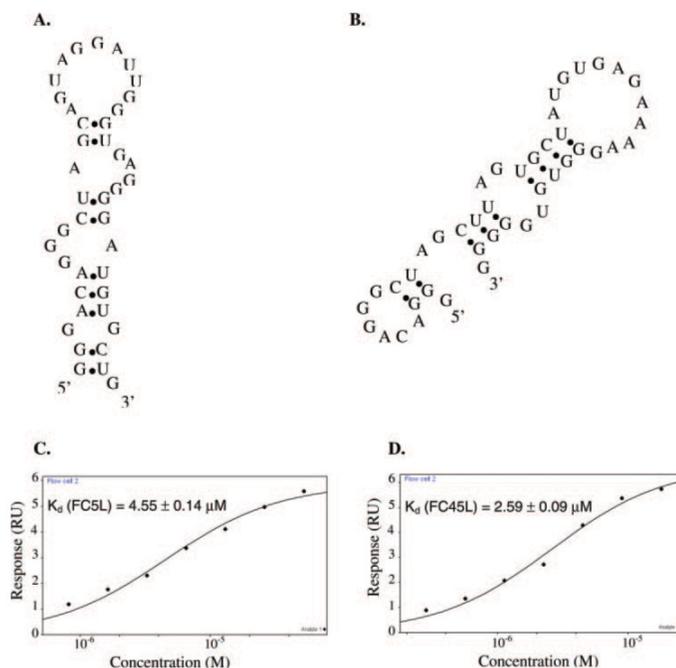
Figure 6. Codeine-binding mini-aptamer characterization. Proposed secondary structures from mfold of (**A**) the FC5 mini-aptamer (FC5L) and (**B**) the FC45 mini-aptamer (FC45L), and the corresponding equilibrium codeine-binding curves of (**C**) FC5L and (**D**) FC45L.

that the modified mini-aptamers bind the target molecule codeine with affinities similar to the corresponding unmodified mini-aptamers. Therefore, these results support the proposed secondary structures of the FC5 and FC45 mini-aptamers (Figure 6) and that their codeine-binding affinities were minimally affected by extending the original base stems. Structural probing studies were performed on FC5 and FC45 full-length aptamers using a standard lead-based cleavage assay to confirm the structures predicted through the SPR analysis. Lead-induced and RNase T1 cleavage patterns were observed to be in agreement with the corresponding proposed structures (Supplementary Figure 4).

The sequence requirements and flexibility of the mini-aptamer base stems were examined with directed mutational analysis coupled with characterization of the effects of these sequence changes on the codeine-binding properties of these aptamers by the described SPR assays. Two of the three original base pairs in the base stem of the FC5 mini-aptamer were replaced with randomly-selected base-pairs (Figure 7C). This new sequence (FC5L-S2) was determined to bind codeine with an affinity ($K_d = 5.39 \pm 0.28$ μM) (Supplementary Figure 2C) comparable to that of the original aptamer sequence, indicating that while the presence of the base stem is essential for codeine-binding, its sequence is not. The sequence space flexibility demonstrated for the aptamer

base stem of FC5L has been reported in other aptamers, such as the theophylline aptamer (33).

Studies were also conducted to demonstrate that the formation of the correct binding pocket within a given aptamer sequence is highly dictated by the formation of the correct base stem. The base stems of two alternative secondary structures for the FC45 mini-aptamer were extended with two GC-pairs to stabilize these proposed secondary structures (Supplementary Figures 3A and B), in the same way as previously described for FC45L-S1. Binding assays revealed that these structures did not bind codeine with as high affinity as the initially proposed structure. The K_d values of these alternative FC45 mini-aptamer structures were increased ∼10-fold (∼25 μM), indicating that the codeine-binding pocket may be somewhat disrupted in these structures. These results indicate that the formation of the correct base stem can have significant influence on the formation of the correct binding pocket for aptamer recognition events.

Validation of the direct coupling SPR assay for characterization of small molecule-aptamer binding properties

Biacore assays are widely used to study a variety of molecular interactions, such as RNA–protein and protein–protein

A. **B.** **C.**
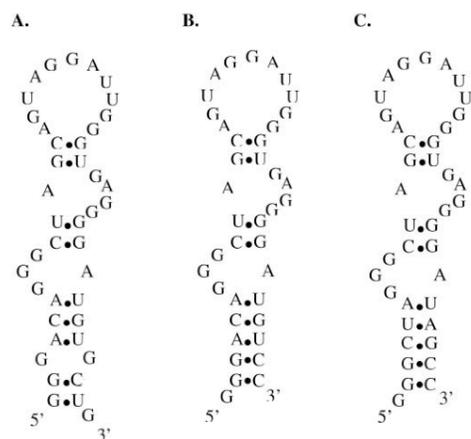
Figure 7. Structural stabilization and sequence requirements of the FC5 mini-aptamer stems. Proposed secondary structures from mfold of (**A**) the original FC5 mini-aptamer (FC5L), (**B**) the FC5 mini-aptamer with a stabilized base stem (FC5L-S1), (**C**) the FC5 mini-aptamer with a stabilized base stem composed of randomly-selected nucleotides (FC5L-S2).

A. **B.**
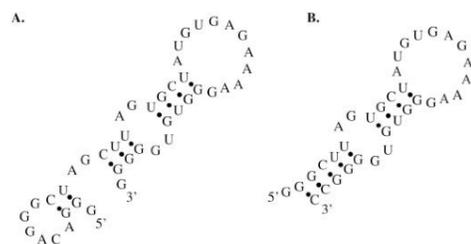
Figure 8. Structural stabilization of the FC45 mini-aptamer. Proposed secondary structures from mfold of (**A**) the original FC45 mini-aptamer (FC45L) and (**B**) the FC45 mini-aptamer with a stabilized base stem (FC45L-S1) in which several nucleotides at the termini of the original mini-aptamer are truncated.

interactions. While these assays are applicable to a broad range of target molecules, proteins have most often been used as the primary targets. Although, Biacore assays have been used to measure the interaction between aptamers and non-protein targets, these assays often include a carrier or linker protein between the target and the sensor surface (24,26). However, significant discrepancies have been observed in K_d values determined from these assays and other commonly used methods potentially due to the use of a linker protein between the dextran surface and the small molecule. While SPR assays in which the small molecule target is directly coupled to the sensor surface without inclusion of a linker protein have been reported previously (22,23), the observed binding properties have not been validated or proven to potentially eliminate non-specific interactions or artifacts

that may arise from the presence of a linker protein used in the assay. Therefore, experiments were conducted to examine the reproducibility, accuracy and versatility of these direct coupling assays.

The reproducibility of the assay method was confirmed through several means. Binding assays were repeated for several samples: two randomly-selected sequences (FC34 and A25), the initial pool, FC5 and FC45. The codeine-binding affinities determined from these replicate experiments were nearly identical, thereby confirming the reproducibility of the assay method (Supplementary Table 1). In addition, the assay was repeated for the initial pool, FC5 and FC45 such that the concentration series sets of these samples were injected into the flow cells in a random order. Consistent K_d values (data not shown) were obtained from the random-injection experiments for all three of the tested RNA samples when compared to the values obtained from injecting them sequentially from lowest to highest concentrations. These results demonstrate the reproducibility and the robustness of this direct coupling assay method. FC45 was used as a positive control when performing the described assays on the remainder of the RNA aptamer sequences.

The potential elimination of non-specific interactions between an aptamer and the linker protein by the direct coupling small molecule-aptamer binding assay was demonstrated on a previously characterized RNA aptamer to a different small molecule target. The described SPR binding assay was performed on a previously characterized dopamine aptamer (dopa2) (18), whose reported K_d value was determined through commonly used solution-based affinity methods. Dopamine was immobilized onto the sensor chip through the same coupling chemistry that was used in the original selection of this dopamine-binding aptamer. The binding affinity determined through the direct coupling SPR assay of the dopa2 RNA aptamer to dopamine ($K_d = 2.71 \pm 0.06$ μM) was nearly identical to the reported value of 2.8 μM (18) (Supplementary Figure 5A).

An indirect coupling assay was performed on the dopamine aptamer using BSA as a protein linker to demonstrate that the presence of a linker protein in a SPR small molecule-aptamer binding assay may generate non-specific interactions or artifacts. It was observed that aptamer samples at the same concentrations take considerably longer to reach an equilibrium binding response in the BSA linker assay versus the direct coupling assay, which is indicative of non-specific interactions. The two highest concentration samples reached a near equilibrium response after 50 min of injection, approximately 33 times longer than that employed in the direct coupling assay. Data analysis revealed that the aptamer binding affinity was significantly affected and resulted in a false assessment as the observed K_d value was substantially higher than the reported value of 2.8 μM (Supplementary Figure 5B). To better analyze the data, Clamp (35) was used to fit the kinetic binding responses, as the two highest concentration samples did not completely reach equilibrium. Kinetic data analysis suggested that multiple binding events are present in the BSA linker assay since the data were well-fit with a multiple binding site model and did not satisfy a one-to-one binding model (Supplementary Figures 5C and D). This finding was further supported by Scatchard plot analysis, which also suggested this indirect coupling assay

as a multiple binding site system represented by a curvature in this plot, a hallmark of a multiple binding site model (Supplementary Figure 5F). In contrast, the direct coupling data fit a single binding site system represented by a linear fit to this data (Supplementary Figure 5E). These results indicate that the presence of a protein linker can cause an aptamer to bind to its surface-immobilized target molecule in a non-specific manner, leading to an inaccurate assessment of the binding affinity of the aptamer to its small molecule target.

DISCUSSION

In this study, we employed *in vitro* selection strategies to isolate RNA aptamers with high affinity and specificity to a subclass of BIA molecules, including codeine, within 15 selection cycles. A counter-selection with morphine and three error-prone PCR steps were incorporated into the selection process to enhance the specificity and affinity of the selected aptamers for their target molecule. The qualitative binding assays revealed that the final aptamer pool was highly enriched with codeine-binding affinity. The binding affinity of the enriched aptamer pool was determined to be 15 μM when characterized through the described Biacore assay; however, several of its member sequences, including FC5 and FC45, were determined to have higher affinities to codeine. In addition both of these aptamers were shown to be highly specific to codeine over morphine, indicating that the morphine counter-selection performed during the selection process was effective at enhancing the desired target specificity of the aptamers. Interestingly, while FC45 maintains codeine-binding specificity over another structural analogue, thebaine, FC5 demonstrates higher specificity to the latter. Therefore, these aptamers exhibit differing specificities to BIA alkaloid molecules, displaying molecular discrimination between targets differing by a single methyl group or structural conformation.

This work also highlights a direct coupling SPR binding assay for accurately and robustly determining the binding properties of aptamers to small molecule ligands. The described method is based on the direct immobilization of the target small molecule onto the sensor chip surface without inclusion of a linker protein as is commonly used. This direct coupling may provide a more accurate assessment of the binding affinity between the small molecule target and the aptamer by eliminating potential non-specific binding between the nucleic acid aptamer and the protein linker and more accurately reproducing conditions used in the selection process. Significant discrepancies have been observed between reported K_d values obtained from Biacore assays that employ a protein linker connecting the target molecule to the sensor surface and other methods that involve direct target coupling. In one example, BSA was used as a linker between a target carbohydrate and the sensor surface (24). The binding affinity of a selected aptamer was reported as 85 pM using this assay method. However, when the aptamer was immobilized onto the sensor surface and target molecules were injected over the surface, the binding affinity to the BSA-linked target was similar to that observed with the earlier experimental setup ($K_d = 57$ pM), whereas the binding affinity to the target molecule alone was determined to be

~60-fold lower ($K_d = 3.3$ nM). In another example, an existing tobramycin aptamer, characterized with a Biacore binding assay using a streptavidin linker, showed a lower degree of selectivity and significantly reduced affinity (25) from previously reported binding properties for this aptamer determined using a number of different assay methods (36–39). These results indicate that the presence of a protein linker may introduce artifacts or non-specificity in the small molecule-aptamer interaction, preventing an accurate assessment of the intact affinity of the aptamer to its target molecule. Direct coupling of the small molecule target onto the sensor surface may provide a more accurate assessment of small molecule-aptamer binding properties by eliminating potential non-specific interactions or artifacts introduced when using a linker protein.

The direct coupling SPR small molecule-aptamer binding assay has the additional benefit of providing a rapid characterization assay. In comparison to other commonly used binding assays, such as isocratic elution or equilibrium filtration, Biacore assays offer a rapid, high-throughput platform, which provides information about both equilibrium and kinetic binding properties. Using the Biacore 2000 and the serial dilution method described in this work, the binding properties of as many as eight aptamer sequences may be accurately and precisely determined in one day on a single chip. It should be noted, that the binding properties determined through this assay correspond to ligand-immobilized binding properties, which may differ from free ligand binding properties depending on the particular aptamer-ligand pair as demonstrated in this and previous work. However, this high-throughput assay strategy may be particularly useful when applied to the screening of libraries for aptamers that exhibit particular binding properties. From this initial screen, those aptamers exhibiting desired binding affinities for surface-immobilized target may be further analyzed with standard solution affinity assays to determine and verify the corresponding binding affinities of those selected aptamers to free target in solution. Furthermore, the high-throughput nature of this platform may be used to rapidly determine the mini-aptamers for selected aptamers through truncation experiments, eliminating the need to perform time-consuming and labor-intensive chemical probing experiments (18,30). In addition, while traditional binding assays involve the use of radiolabeled aptamers or often rare and expensive radiolabeled target molecules, Biacore assays eliminate this requirement. The same assay methodology may be employed to perform structural stabilization studies, which were used to develop mini-aptamers with stabilized base stems. Aptamers with stabilized, modifiable and extendable base stems are more functionally attractive for applications in downstream molecular design strategies that involve exploiting structural rearrangements associated with the base stem formation (17). Therefore, the FC5 and FC45 mini-aptamers may be readily employed in molecular engineering applications as their stems are extendable and modifiable. Finally, the versatility of this direct coupling SPR assay to the study of small molecule-aptamer interactions was demonstrated through several means and validated on a previously characterized dopamine RNA aptamer. Elimination of non-specific interactions was demonstrated in the direct coupling assay compared to the indirect coupling assay for the same aptamer, where

non-specific interactions or binding artifacts arose in the presence of the linker protein. Therefore, the SPR assay discussed here is proven to be a rapid, versatile, accurate and robust method for quantitative measurement of small molecule–RNA interactions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. William,D.G., Hatch,D.J. and Howard,R.F. (2001) Codeine phosphate in paediatric medicine. *Br. J. Anaesth.*, **86**, 413–421.
2. Williams,D.G., Patel,A. and Howard,R.F. (2002) Pharmacogenetics of codeine metabolism in an urban population of children and its implications for analgesic reliability. *Br. J. Anaesth.*, **89**, 839–845.
3. Kim,I., Barnes,A.J., Oyler,J.M., Schepers,R., Joseph,R.E., Jr, Cone,E.J., Lafko,D., Moolchan,E.T. and Huestis,M.A. (2002) Plasma and oral fluid pharmacokinetics and pharmacodynamics after oral codeine administration. *Clin. Chem.*, **48**, 1486–1496.
4. Jayasena,S.D. (1999) Aptamers: an emerging class of molecules that rival antibodies in diagnostics. *Clin. Chem.*, **45**, 1628–1650.
5. Silverman,S.K. (2003) Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA *RNA*, **9**, 377–383.
6. O'Sullivan,C.K. (2002) Aptasensors--the future of biosensing? *Anal. Bioanal. Chem.*, **372**, 44–48.
7. Hesselberth,J.R., Robertson,M.P., Knudsen,S.M. and Ellington,A.D. (2003) Simultaneous detection of diverse analytes with an aptazyme ligase array. *Anal. Biochem.*, **312**, 106–112.
8. Seetharaman,S., Zivarts,M., Sudarsan,N. and Breaker,R.R. (2001) Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nat. Biotechnol.*, **19**, 336–341.
9. Facchini,P.J. (2001) Alkaloid biosynthesis in plants: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **52**, 29–66.
10. Allen,R.S., Millgate,A.G., Chitty,J.A., Thisleton,J., Miller,J.A., Fist,A.J., Gerlach,W.L. and Larkin,P.J. (2004) RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat. Biotechnol.*, **22**, 1559–1566.
11. Rathbone,D.A. and Bruce,N.C. (2002) Microbial transformation of alkaloids. *Curr. Opin. Microbiol.*, **5**, 274–281.
12. Suess,B., Hanson,S., Berens,C., Fink,B., Schroeder,R. and Hillen,W. (2003) Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res.*, **31**, 1853–1858.
13. Suess,B., Fink,B., Berens,C., Stentz,R. and Hillen,W. (2004) A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*. *Nucleic Acids Res.*, **32**, 1610–1614.
14. Werstuck,G. and Green,M.R. (1998) Controlling gene expression in living cells through small molecule–RNA interactions. *Science*, **282**, 296–298.
15. Desai,S.K. and Gallivan,J.P. (2004) Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.*, **126**, 13247–13254.
16. Winkler,W.C. and Breaker,R.R. (2005) Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.*, **59**, 487–517.
17. Bayer,T.S. and Smolke,C.D. (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat. Biotechnol.*, **23**, 337–343.
18. Mannironi,C., Di Nardo,A., Fruscoloni,P. and Tocchini-Valentini,G.P. (1997) *In vitro* selection of dopamine RNA ligands. *Biochemistry*, **36**, 9726–9734.
19. Mannironi,C., Scerch,C., Fruscoloni,P. and Tocchini-Valentini,G.P. (2000) Molecular recognition of amino acids by RNA aptamers: the evolution into an L-tyrosine binder of a dopamine-binding RNA motif. *RNA*, **6**, 520–527.
20. Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.
21. Ellington,A.D. and Szostak,J.W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.
22. Kwon,M., Chun,S.M., Jeong,S. and Yu,J. (2001) *In vitro* selection of RNA against kanamycin B. *Mol. Cells*, **11**, 303–311.
23. Gebhardt,K., Shokraei,A., Babaie,E. and Lindqvist,B.H. (2000) RNA aptamers to S-adenosylhomocysteine: kinetic properties, divalent cation dependency, and comparison with anti-S-adenosylhomocysteine antibody. *Biochemistry*, **39**, 7255–7265.
24. Jeong,S., Eom,T., Kim,S., Lee,S. and Yu,J. (2001) *In vitro* selection of the RNA aptamer against the Sialyl Lewis X and its inhibition of the cell adhesion. *Biochem. Biophys. Res. Commun.*, **281**, 237–243.
25. Verhelst,S.H., Michiels,P.J., van der Marel,G.A., van Boeckel,C.A. and van Boom,J.H. (2004) Surface plasmon resonance evaluation of various aminoglycoside-RNA hairpin interactions reveals low degree of selectivity. *Chembiochem*, **5**, 937–942.
26. Davis,J.H. and Szostak,J.W. (2002) Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc. Natl Acad. Sci. USA*, **99**, 11616–11621.
27. Arnold,F.H. and Blanch,H.W. (1986) Analytical affinity chromatography II. Rate theory and the measurement of biological binding kinetics. *J. Chromatogr.*, **355**, 13–27.
28. Brockstedt,U., Uzarowska,A., Montpetit,A., Pfau,W. and Labuda,D. (2004) *In vitro* evolution of RNA aptamers recognizing carcinogenic aromatic amines. *Biochem. Biophys. Res. Commun.*, **313**, 1004–1008.
29. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
30. Berens,C., Thain,A. and Schroeder,R. (2001) A tetracycline-binding RNA aptamer. *Bioorg. Med. Chem.*, **9**, 2549–2556.
31. Lozupone,C., Changayil,S., Majerfeld,I. and Yarus,M. (2003) Selection of the simplest RNA that binds isoleucine. *RNA*, **9**, 1315–1322.
32. Majerfeld,I. and Yarus,M. (1998) Isoleucine:RNA sites with associated coding sequences. *RNA*, **4**, 471–478.
33. Jenison,R.D., Gill,S.C., Pardi,A. and Polisky,B. (1994) High-resolution molecular discrimination by RNA. *Science*, **263**, 1425–1429.
34. Haller,A.A. and Sarnow,P. (1997) *In vitro* selection of a 7-methyl-guanosine binding RNA that inhibits translation of capped mRNA molecules. *Proc. Natl Acad. Sci. USA*, **94**, 8521–8526.
35. Myszka,D.G. and Morton,T.A. (1998) CLAMP: a biosensor kinetic data analysis program. *Trends Biochem. Sci.*, **23**, 149–150.
36. Wang,Y. and Rando,R.R. (1995) Specific binding of aminoglycoside antibiotics to RNA. *Chem. Biol.*, **2**, 281–290.
37. Wang,Y., Killian,J., Hamasaki,K. and Rando,R.R. (1996) RNA molecules that specifically and stoichiometrically bind aminoglycoside antibiotics with high affinities. *Biochemistry*, **35**, 12338–12346.
38. Hamasaki,K., Killian,J., Cho,J. and Rando,R.R. (1998) Minimal RNA constructs that specifically bind aminoglycoside antibiotics with high affinities. *Biochemistry*, **37**, 656–663.
39. Cho,J., Hamasaki,K. and Rando,R.R. (1998) The binding site of a specific aminoglycoside binding RNA molecule. *Biochemistry*, **37**, 4985–4992.

APPENDIX F

Vol. 83, No. 1

Design and Expression of a Dimeric Form of Human Immunodeficiency Virus Type 1 Antibody 2G12 with Increased Neutralization Potency[▽]

Anthony P. West, Jr.,[1] Rachel P. Galimidi,[1] Christopher P. Foglesong,[1] Priyanthi N. P. Gnanapragasam,[1] Kathryn E. Huey-Tubman,[1,2] Joshua S. Klein,[1] Maria D. Suzuki,[1] Noreen E. Tiangco,[1] Jost Vielmetter,[1,3] and Pamela J. Bjorkman[1,2]*

Division of Biology,[1] Howard Hughes Medical Institute,[2] and Caltech Protein Expression Center,[3] California Institute of Technology, 1200 E. California Blvd., Pasadena, California 91125

The antigen-binding fragment of the broadly neutralizing human immunodeficiency virus type 1 (HIV-1) antibody 2G12 has an unusual three-dimensional (3D) domain-swapped structure with two aligned combining sites that facilitates recognition of its carbohydrate epitope on gp120. When expressed as an intact immunoglobulin G (IgG), 2G12 formed typical IgG monomers containing two combining sites and a small fraction of a higher-molecular-weight species, which showed a significant increase in neutralization potency (50- to 80-fold compared to 2G12 monomer) across a range of clade A and B strains of HIV-1. Here we show that the higher-molecular-weight species corresponds to a 2G12 dimer containing four combining sites and present a model for how intermolecular 3D domain swapping could create a 2G12 dimer. Based on the structural model for a 3D domain-swapped 2G12 dimer, we designed and tested a series of 2G12 mutants predicted to increase the ratio of 2G12 dimer to monomer. We report a mutation that effectively increases the 2G12 dimer/monomer ratio without decreasing the expression yield. Increasing the proportion of 2G12 dimer compared to monomer could lead to a more potent reagent for gene therapy or passive immunization.

Broadly neutralizing antibodies against human immunodeficiency virus type 1 (HIV-1) have attracted attention not only for the lessons they provide for designing vaccine antigens to induce a more robust immunological response (2) but also as potential therapeutic reagents. Although HIV infection leads to a vigorous antibody response, most antibodies fail to control the virus due to targeting of non-neutralizing epitopes or the ability of escape mutants to quickly develop against neutralizing antibodies (23). Correlating with the ability of the virus to elude antibodies, the majority of neutralizing antibodies are highly strain specific. Nevertheless, a small set of broadly neutralizing antibodies has been isolated from the blood of HIV-infected individuals, and these reagents have been extensively studied (2). Clinical trials using a cocktail of three such antibodies—2G12, 4E10, and 2F5—have demonstrated a partial ability to suppress viral replication (13, 20, 21).

The 2G12 antibody has an unusual structure that facilitates recognition of its carbohydrate epitope on gp120 (4). Whereas typical immunoglobulin G (IgG) antibodies contain two flexibly attached antigen-binding fragments (Fabs), resulting in two antigen-binding sites separated by distances ranging from 120 to 150 Å in structures of intact IgGs (6, 7, 17), the Fab arms of 2G12 are entwined in such a way as to create a single antigen-binding region with two rigidly arranged antigen-binding sites separated by ~35 Å (4) (Fig. 1A and B). The entwined structure of the 2G12 Fabs results from three-dimensional (3D) domain swapping (1) in which each 2G12 light chain associates with both heavy chains: the light-chain variable domain (V_L) is paired with the variable domain of one heavy chain (V_H), while the light constant domain (C_L) is paired with constant domain 1 (C_H1) of the partner heavy chain (Fig. 1B). This domain-swapped arrangement prevents the Fab arms from having the normal flexibility observed in other antibodies but, by possessing a double-sized antigen-combining site, the 2G12 Fab_2 unit is able to recognize clusters of mannose-rich carbohydrates that occur on gp120 (18). Normally, these carbohydrates create a glycan shield on the HIV envelope glycoprotein (Env) spike that helps the virus evade the host antibody response (23).

During expression of 2G12 in mammalian cells, we observed the production of IgG monomers (i.e., two heavy chains, two light chains, and thus two Fabs), as typically formed by other IgGs, and a higher-molecular-weight fraction that exhibited a significantly increased neutralization potency. Here, we show that the higher-molecular-weight fraction corresponded to a 2G12 dimer with four heavy chains, four light chains, and four Fabs and present a model for how 3D domain swapping could create a 2G12 dimer. We used the model of the 2G12 dimer to design mutations predicted to increase the fraction of dimer being expressed and report a mutation that effectively increases the 2G12 dimer/monomer ratio without decreasing the expression yield.

MATERIALS AND METHODS

Materials. Sequences encoding the 2G12 V_H-C_H1 and V_L-C_L domains in the pComb3H expression vector (a gift from Dennis Burton, The Scripps Research Institute) were subcloned into the bicistronic baculovirus vector pAc-κ-Fc (PROGEN Biotechnik), which contains the gene for the Fc region of human IgG1 (G1m marker). The heavy-chain gene, now including the hinge and Fc regions (residues 236 to 446, numbered according to the method of Kabat et al. [11]), and light-chain gene were each subcloned into the mammalian expression vector pTT5 (NRC Biotechnology Research Institute) for expression in

* Corresponding author. Mailing address: Division of Biology, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125. Phone: (626) 395-8350. Fax: (626) 395-5939. E-mail: bjorkman@caltech.edu.
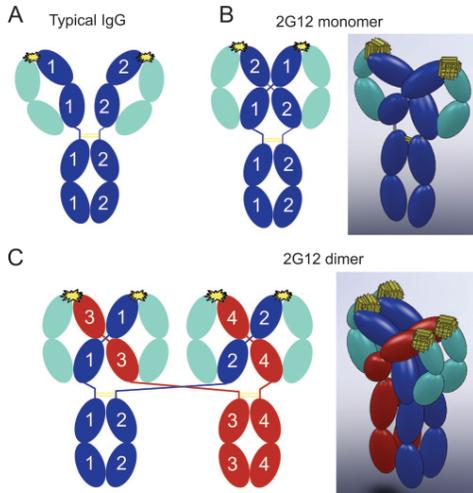[▽] Published ahead of print on 22 October 2008.

FIG. 1. Schematic structures of a typical IgG and 2G12. Heavy chains are blue in panels A and B and blue or red in panel C, light chains are cyan, disulfide bonds are yellow lines, and the antigen combining sites are yellow starbursts. (A) Schematic diagram showing the domain arrangement of a typical IgG, which contains two identical heavy chains and two identical light chains. (B) Schematic diagram (left) and a corresponding 3D model (right) illustrating chain pairing in monomeric 2G12 (based on structural data from reference 4). As a result of intramolecular 3D domain swapping, each heavy chain forms part of both Fab units to create a rigidly arranged Fab$_2$ unit. To distinguish the two heavy chains, they are labeled 1 or 2 in the schematic diagram. (C) Schematic diagram (left) and corresponding 3D model (right) illustrating chain pairing in dimeric 2G12. The proposed dimer structure resulting from intermolecular 3D domains swapping has the same domain-swapped Fab$_2$ unit as the monomer, but the connectivity to the Fc domains is altered. To distinguish the four heavy chains, they are labeled 1, 2, 3, or 4 in the schematic diagram and are red in one of the IgG monomer precursors.

HEK293-6E cells. Mutations were introduced in the 2G12 heavy-chain gene by using a QuikChange mutagenesis kit (Stratagene). Seven different forms of the 2G12 heavy-chain gene were constructed: our original 2G12 sequence, which differed from previously reported 2G12 and IgG1 Fc sequences (4) by two substitutions (V5L and H237S), and six hinge deletion mutants constructed from our original 2G12 sequence. The mutants were D1 (deletion of residue 237), D2 (deletion of residues 236 to 237), D3 (deletion of residues 235 to 237), D4 (deletion of residues 232 to 237), D6 (deletion of residues 232 to 239), and D6GG (deletion of residues 232 to 239 and two proline-to-glycine substitutions [P240G and P241G]) (see Fig. 3A). In addition, we produced forms of the original 2G12 sequence and of the D2 mutant that reverted the V5L and H237S substitutions to the previously reported 2G12 and IgG1 Fc sequences.

Expression of 2G12. Proteins were expressed by using polyethylenimine-mediated transient transfection (5) of suspension HEK293-6E cells (NRC Biotechnology Research Institute). The 2G12 heavy- and light-chain expression vectors were cotransfected at a 1:1 ratio using 25-kDa linear polyethylenimine (Polysciences). Cell culture supernatants were collected at 6 days posttransfection, passed over protein A resin (Pierce Biotechnology), and eluted using a pH 3.0 citrate buffer. Eluates from the protein A column were immediately neutralized and then subjected to size exclusion chromatography in 20 mM Tris (pH 8.0)–150 mM NaCl using a Superdex 200 16/60 or 10/30 column (GE Healthcare). The column profile revealed two major peaks, later identified as 2G12 dimer and 2G12 monomer. Similar results were obtained when the original 2G12 construct and the D2 mutant were purified from cell culture supernatants by passage over

a neonatal Fc receptor affinity column in which supernatants were passed over the column at pH 6.0 and eluted at pH 8.0 (8).

Since the sizing column profiles of the monomeric and dimeric forms of 2G12 partially overlapped, samples for neutralization assays were subjected to two size exclusion chromatography purification steps: first, the concentrated protein A eluate was run over a Superdex 200 16/60 column, and then the monomer and dimer fractions were placed over a Superdex 200 10/30 column.

Parallel trials of the original 2G12 construct and six hinge deletion mutants were performed on a 400-ml scale. Five independent expression trials, each involving all seven constructs, were conducted. In each trial, 2G12 proteins were purified on a protein A column as described above and then subjected to size exclusion chromatography. Dimer-to-monomer ratios for each construct (see Fig. 3C) were calculated after integration of the peak areas using UNICORN software. In four of the five trials, the expression of total 2G12 protein in each sample was approximately 2 mg/liter. In the fifth trial, the total expression in each sample was ~10-fold lower, likely due to suboptimal transfections. The data from this trial were omitted from the calculated means shown in Fig. 3C.

Static light scattering. The oligomeric states of the two 2G12 peaks isolated by conventional size exclusion chromatography were determined using size exclusion chromatography with in-line static light scattering and refractive index monitoring. Experiments were performed at 25°C using an ÄKTA chromatography system (GE Healthcare) with a Superdex 200 10/30 (GE Healthcare) equipped with a Dawn Helios multi-angle light scattering photometer and an Optilab rEX refractive index detector (Wyatt Technology). Bovine serum albumin was used as a calibration standard. All data were analyzed with ASTRA V software (Wyatt Technology).

In vitro neutralization assays. We used a previously described pseudovirus neutralization assay, which measures the reduction in luciferase reporter gene expression in the presence of a potential inhibitor following a single round of pseudovirus infection in TZM-bl cells (12, 14). Pseudoviruses were generated by cotransfection of HEK293T cells with an Env expression plasmid and a replication-defective backbone plasmid. Neutralization assays were performed either in-house (data shown only for strain HxBc2) or by the Collaboration for AIDS Vaccine Discovery (CAVD) core neutralization facility (Table 1). Strains SC422661.8, TRO.11, PVO.4, QH0692.42, 6535.3, and HxBc2 were tested in-house, with results similar to those obtained from the CAVD assays. For in-house assays, each 2G12 sample was tested in triplicate, with 200 infectious viral units per well incubated with a threefold dilution series. After a 1-h incubation at 37°C, 10,000 TZM-bl cells were added to each well, followed by incubation for 2 days. Cells were then lysed and assayed for luciferase expression by using Bright-Glo (Promega) and a Victor3 luminometer (Perkin-Elmer). The percent neutralization was determined by calculating the difference in luminescence between test wells and cell control wells (cell only), dividing this value by the difference between virus control (cell plus virus) and cell control wells, subtracting from 1, and multiplying by 100. Nonlinear regression analysis was used to calculate concentrations at which half-maximal inhibition was observed (IC$_{50}$s). The average IC$_{50}$s reported in Table 1 are geometric means calculated by using the formula $(\Pi a_i)^{(1/n)}$, where $i = 1, 2, \ldots, n$. Calculation of geometric means is suitable for data sets covering multiple orders of magnitude (19), as is the case for neutralization data across multiple viral strains.

Biacore binding studies. A Biacore T100 biosensor system (GE Healthcare) was used to evaluate binding of the 2G12 proteins to gp120. In this system, a protein is coupled to a gold-dextran layer, and association and dissociation phases for binding to an injected protein are measured in real time in resonance units (RU). The original 2G12 monomer and dimer and the D2 dimer (~150 RU of each) were captured onto ~4,000 RU of goat anti-Fc polyclonal antibody (Chromapure; Jackson ImmunoResearch), which was immobilized by primary amine coupling to a CM5 sensor chip as described in the Biacore manual. A concentration series of monomeric gp120 (expressed in baculovirus-infected insect cells; strain HxBc2) was injected at 20 µl/min over the flow cells. After the gp120 dissociation phase, the surface was regenerated by injection of pH 1.5 glycine buffer, and 2G12 proteins were captured again prior to the subsequent gp120 injection.

Structural analysis. The 2G12 structure was displayed and analyzed by using the program O (10). The 3D model of 2G12 was prepared using SolidWorks (SolidWorks Corp.).

RESULTS

Identification of a dimeric form of 2G12 with increased neutralization potency. During purification of 2G12 IgG expressed in mammalian cells, we noticed that the protein eluted

TABLE 1. IC_{50}s for four forms of 2G12 obtained from in vitro pseudovirus neutralization assays of 2G12-sensitive strains[a]

Env clone	Clade	2G12 original			D2		
		Monomer IC_{50} (µg/ml)	Dimer IC_{50} (µg/ml)	Monomer IC_{50}/ dimer IC_{50}	Monomer IC_{50} (µg/ml)	Dimer IC_{50} (µg/ml)	Monomer IC_{50}/ dimer IC_{50}
6535.3	B	30.1	0.34	89	32.9	1.5	22
QH0692.42	B	9.8	0.08	123	11.1	0.1	111
SC422661.8	B	17.8	<0.05	356	19.4	0.1	194
PVO.4	B	4.9	<0.05	98	7.1	<0.05	142
TRO.11	B	1.6	<0.05	32	1.4	<0.05	28
WITO4160.33	B	6.8	<0.05	136	7.6	0.05	152
3988.25	B	3.2	<0.05	64	2.8	<0.05	56
7165.18	B	6.6	<0.05	132	8	<0.05	160
QH0515.1	B	0.1	<0.05	2	0.1	<0.05	2
5768.4	B	0.1	<0.05	2	0.1	<0.05	2
6101.1	B	24	0.05	480	26.5	0.06	442
TRJO4551.58	B	>100	0.16	625	ND[b]	ND	ND
HxBc2*	B	0.47	0.028	17	ND	ND	ND
0330.v4.c3	A	15.3	<0.05	306	17.2	<0.05	344
3415.v1.c1	A	8	<0.05	160	8.1	<0.05	162
CHO38.12	CRF07_BC	0.2	<0.05	4	0.3	<0.05	6
211-9	CRF02_AG	>100	0.4	250	>100	1.3	77
235-47	CRF02_AG	1.3	<0.05	26	1.4	<0.05	28
T280-5	CRF02_AG	>100	1	100	51	4.9	10
T250-4	CRF02_AG	61.5	<0.05	1230	60.7	0.4	152
T251	CRF02_AG	50.8	0.1	508	46.6	0.3	155
Avg (geometric mean)		6.3	0.08	82	6.5	0.12	54

[a] Results from in-house neutralization assays are indicated by an asterisk; all other results were obtained by the CAVD core neutralization facility. In calculating average IC_{50}s, measurements outside of the range of the assay (<0.05 or >100 µg/ml) were assigned to those limiting values, and averages were calculated using only strains for which IC_{50}s were available for all four forms of 2G12. The following clade B and C strains were not neutralized by the original 2G12 monomer or dimer (IC_{50} values > 100 µg/ml): AC10.0.29, RHPA4259.7, THRO4156.18, REJO4541.67, and CAAN5342.A2 (clade B) and Du156.12, Du172.17, Du422.1, ZM197M.PB7, ZM214M.PL15, ZM233M.PB6, ZM249M.PL1, ZM53M.PB12, ZM109F.PB4, ZM135M.PL10a, CAP45.2.00.G3, and CAP210.2.00.E8 (clade C).

[b] ND, not determined.

as two peaks from a size exclusion chromatography column (peak 1, eluting at ~12 ml, and peak 2, eluting at ~10 ml) (Fig. 2A). Peak 1 migrated as if it were a somewhat larger protein than a typical IgG, which eluted at ~13 ml, suggesting that it corresponded to a 2G12 monomer, which has a more elongated structure than other IgG monomers due to the 3D domain-swapped structure of its Fabs (4) (Fig. 1B). Peak 2 appeared to correspond to a higher-molecular-weight form of 2G12. Samples from both peaks migrated on sodium dodecyl sulfate-polyacrylamide gel electrophoresis under nonreducing conditions, as expected for an IgG (Fig. 2B), indicating that covalent modification(s) could not account for the altered migration of peak 2. Once isolated, both peaks retained their size exclusion profile over a period of weeks to months, indicating that there was little tendency for conversion between the two forms.

In-line static multi-angle light scattering was used to determine the absolute molecular masses, and therefore the oligomeric states, of both peaks. For this analysis, we used the major portion of peak 2 and did not include the shoulder slightly ahead of the peak. The experimentally determined molecular masses derived from these data were 135 kDa for peak 1 and 285 kDa for peak 2. By comparing these values with the molecular mass predicted from the 2G12 sequence (145 kDa), we identified peak 1 as 2G12 monomer and peak 2 as 2G12 dimer.

We next compared the potencies of purified 2G12 monomer and dimer to neutralize a range of HIV-1 pseudoviruses by using an in vitro neutralization assay (12, 14). The 2G12 mono-

mer neutralized a range of clade A and B, but not clade C, strains (Table 1). Over the range of strains that were sensitive to 2G12 neutralization, the 2G12 dimer was generally 1 to 2 orders of magnitude more potent than the 2G12 monomer, with an average increased potency of 82-fold (Table 1). Three of the twenty strains that were not neutralized by monomeric 2G12 (IC_{50} > 100 µg/ml) were neutralized by dimeric 2G12: strains TRJO4551.58 (clade B) and strains 211-9 and T280-5 (circulating recombinant forms) (Table 1). The clade C strains remained resistant to 2G12 dimer (Table 1 legend).

Design and expression of 2G12 mutants. Given that the 2G12 dimer was not formed by covalent joining of two monomers (Fig. 2B), we reasoned that the dimeric form could arise from intermolecular 3D domain swapping in which each light chain was paired with heavy chains derived from two IgGs (Fig. 1C). In contrast, the intramolecular domain swapping observed in monomeric 2G12 involves the pairing of each light chain with both heavy chains from a single IgG (4) (Fig. 1B).

To increase the dimeric fraction of 2G12, we designed mutants predicted to favor intermolecular domain swapping instead of the intramolecular swapping that led to the monomer. Compared to a typical antibody, the hinge regions of the domain-swapped 2G12 are forced into a relatively extended conformation. In other words, the C-terminal ends of the 2G12 C_H1 domains (defined by residue K228) are held fixed at a large distance: the K228-K228 distance is 40.6 Å in the 2G12 Fab_2 structure (pdb code 1OP3) (4) versus 25.0 Å for the corresponding distance in an unswapped antibody structure
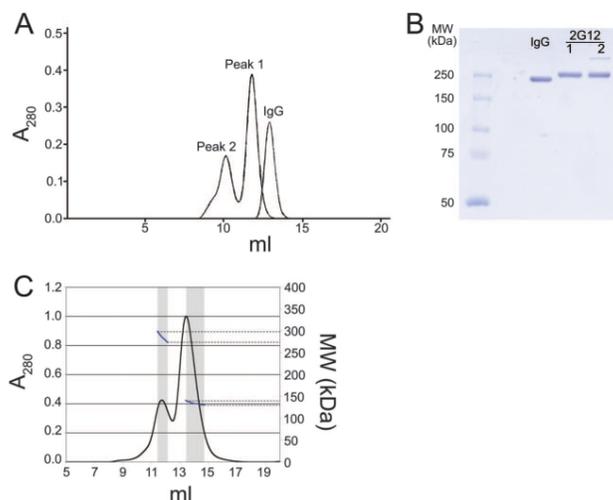
FIG. 2. Size exclusion chromatography profile of 2G12. (A) Comparison of the elution profiles of protein A-purified 2G12 with a typical human IgG. The more elongated structure of the 2G12 IgG monomer (peak 1) caused it to elute slightly earlier than typical IgGs. (B) Sodium dodecyl sulfate–7.5% polyacrylamide gel electrophoresis of a typical human IgG and the 2G12 proteins in peaks 1 and 2. Samples were run under nonreducing conditions. (C) Multi-angle light scattering data obtained by size exclusion chromatography with in-line light scattering and refractive index monitoring. UV absorbance at 280 nm is shown in black, and the calculated molecular weight based on multi-angle light scattering data is indicated in blue with units shown on the right axis. The windows used for calculating the molecular masses of peak 1 and peak 2 are shown in gray.

(b12 IgG; pdb code 1HZH) (17). Intramolecular domain swapping, as occurs in monomeric 2G12, may be disfavored by a shorter hinge region, which could instead accommodate an intermolecular domain swap resulting in an intertwined IgG dimer instead of an intertwined IgG monomer (Fig. 1B and C). Based on this structural hypothesis we designed forms of 2G12 with shorter hinge regions.

A series of deletions was made in the 2G12 heavy chain between residue C230, which normally participates in a disulfide bond with the light chain, and residues C239 and C242, which form disulfide bonds between the heavy chains (Fig. 3A). The mutants were named according to how many residues were deleted, with D1 referring to a single residue deletion, up to D6, which represented a six-residue deletion. An additional mutant, D6GG, in which two prolines were replaced with glycines to allow increased flexibility, was also prepared.

The original 2G12 construct was expressed, along with the suite of mutants in multiple expression trials. During each trial, the expressed IgGs were purified from transfected cell supernatants on a protein A column, followed by size exclusion chromatography to separate 2G12 monomer and dimer. The amounts of dimer and monomer produced by each construct were determined by integrating the area under the relevant size exclusion chromatography peak (Fig. 3B). As shown in Fig. 3C, the D2 mutant had a significantly higher dimer/monomer ratio (0.66 ± 0.06) than the original 2G12 (0.28 ± 0.03). The other mutants showed only slightly higher dimer/monomer ratios (D3 to D6, D6GG) or a lower ratio (D1) than the original 2G12. A comparison of the total expression levels

(monomer plus dimer) of the original and D2 constructs showed that the D2 mutation did not impair overall expression (Fig. 3C).

2G12 protein has been reported to aggregate under strongly acidic conditions (16). To address whether the low pH used during elution from the protein A column affected the ability of 2G12 to dimerize, we purified 2G12 from transfected cell supernatants by using a neonatal Fc receptor affinity column in which samples were loaded at pH 6.0 and eluted at pH 8.0 (8). Size exclusion chromatography profiles of the original 2G12 and the D2 mutant were similar for both purification schemes, indicating that dimer formation was not affected by the protein A elution conditions. In addition, we verified that our results were not influenced by two substitutions in our 2G12 construct compared to the published sequences of 2G12 Fab_2 and human IgG1 Fc that were introduced during cloning (see Materials and Methods). To ascertain the effects of the substitutions, we produced versions of the original 2G12 construct and the D2 mutant in which the substitutions were reverted and found no significant differences in the total yields of 2G12 protein and the dimer/monomer ratios for the reverted constructs compared to the constructs containing the two substitutions: the average total yields for both reverted constructs were within 8% of the unreverted constructs, and the dimer/monomer ratios for the reverted 2G12 and the D2 mutant constructs were 0.34 ± 0.03 (reverted 2G12) and 0.76 ± 0.06 (reverted D2 mutant).

To verify that the substitutions in the D2 mutant did not affect its neutralization activity, we determined the neutraliza-
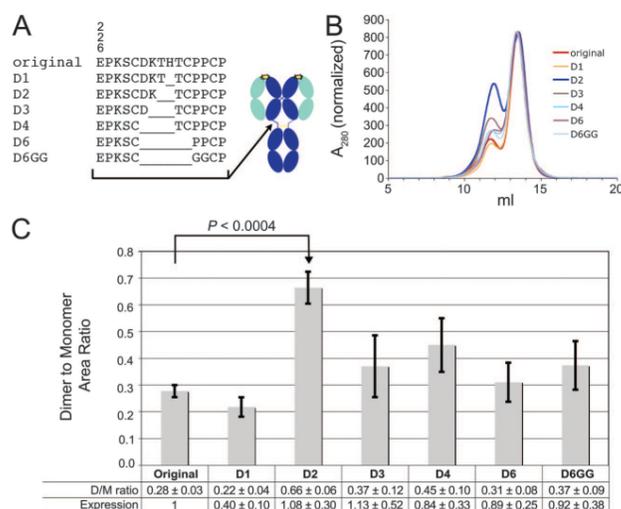
FIG. 3. 2G12 deletion constructs. (A) Amino acid sequences of the hinge region of the original 2G12 and the six deletion mutants (left) and approximate location of the hinge deletions on a schematic structure of a normal IgG (right). (B) Size exclusion chromatography profiles of the original 2G12 and the deletion mutants. Absorbance traces are shown for one expression trial, normalized on the monomer peak height. (C) Dimer/monomer ratios for the original 2G12 and the deletion mutants. The data are presented as the means and standard deviations for ratios determined from four independent expression trials, each of which involved the expression of all seven constructs. Also shown is the mean total expression relative to the original 2G12 construct. The P value comparing the original 2G12 and the D2 mutant was computed by using the two-sided independent sample Student t test.

tion potencies of monomeric and dimeric D2 for the strains that were tested for the original 2G12 monomer and dimer. As seen for the original 2G12 dimer, the D2 dimer was more potent in neutralization of 2G12-sensitive strains than the corresponding monomer, and two of the strains that were resistant to the original 2G12 monomers were sensitive to D2 2G12 dimers (Table 1). The average IC_{50} for the D2 dimer (0.12 μg/ml) was slightly higher than the average value for the original dimer (0.08 μg/ml), and the overall monomer IC_{50}/dimer IC_{50} ratio for the D2 mutant (54-fold) was lower than the corresponding ratio for the original 2G12 dimer (82-fold). However, the differences in neutralization potencies of the D2 and original 2G12 dimers are not necessarily significant because the dimer IC_{50} averages that were used to derive the ratios were calculated using strains with IC_{50}s that were below the sensitivity limit (0.05 μg/ml) of our assay; these are reported as <0.05 μg/ml in Table 1 and input at 0.05 μg/ml in the calculations of averages.

To address whether dimeric 2G12 exhibits higher neutralization potencies than monomeric 2G12 because of avidity effects (i.e., higher apparent affinity resulting from multivalent binding), we compared the binding of gp120 to the two forms of 2G12. In these experiments, the original 2G12 monomer and dimer and the D2 dimer were coupled to a biosensor chip, and the binding of injected monomeric HxBc2 gp120 was evaluated. No significant differences were observed for the interactions with gp120 of the dimers compared to the monomer: the equilibrium dissociation constants (K_Ds) were ~1 nM for

the binding of gp120 to the original monomer, the original dimer, and the D2 dimer (data not shown). These results do not rule out avidity effects as a mechanism for increased neutralization potency, however, since the gp120 monomers would not contain the same arrangement of carbohydrate epitopes as a trimeric gp120/gp41 envelope spike on the surface of a virus.

DISCUSSION

Unlike other known broadly neutralizing anti-HIV-1 antibodies, 2G12 recognizes a carbohydrate, rather than protein, epitope on gp120 (18, 22). The unusual domain-swapped structure of its Fab_2 unit facilitates recognition of adjacent oligomannose units on the same envelope spike (3). Multimerization of 2G12 has been reported to increase the neutralization potency of 2G12; an engineered IgM form of 2G12, which contained 10 Fabs (as compared to 2 in a monomeric IgG), was shown to neutralize up to 28-fold better than monomeric 2G12 IgG (24). Previous investigators suggested that 2G12 IgG might form dimers with higher neutralization potency; for example, the severalfold higher neutralization potency for 2G12 produced in transgenic maize as opposed to mammalian cells was ascribed to a higher dimer or aggregate content (15), and 2G12 IgG dimers are mentioned but not characterized or directly compared to 2G12 IgM or 2G12 monomers by Wolbank et al. (24). Here we report a higher-molecular-weight form of 2G12 IgG that is naturally produced during expression in

mammalian cells, and we present biophysical evidence that it represents a dimeric IgG.

We compared the potencies of isolated 2G12 dimer and monomer for neutralization of a range of clade A and clade B HIV-1 strains, and report that the 2G12 dimer has a >50-fold average increased potency compared to the monomer. The high neutralization potency of 2G12 dimers suggests that the dimeric fraction of 2G12 would be a better candidate for passive immunotherapy to treat HIV-1 infection than the monomeric fraction, which appears to have been used in previous studies of the efficacy of passive immunization using anti-HIV antibodies (13, 21). Another form of 2G12 that could be considered for passive immunizations is an engineered IgM form, which was reported to exhibit up to an ~28-fold increase in neutralization potency compared to 2G12 IgG (24). Because the previous study used a different neutralization assay and different viral strains, the average neutralization potency increases cannot be directly compared, but their similarity in magnitude suggests that the extra Fab$_2$ units in the pentameric IgM form compared to the dimeric form do not contribute significantly to more effective neutralization. In any case, the choice of an optimal therapeutic reagent depends on other factors in addition to neutralization potency, including in vivo half-life, tissue distribution, antibody-dependent cellular cytotoxicity activity, and expression yield. Given the longer serum half-lives of IgG versus IgM antibodies (9), the activity of IgG antibodies in antibody-dependent cellular cytotoxicity (9) and the more complex assembly and heterogeneity of recombinant J chain-containing IgM and IgA, the dimeric IgG form of 2G12 is likely to have several advantages compared to the IgM form.

Based on the propensity of 2G12 monomers to form an intramolecular domain-swapped Fab$_2$ structure, we present a structural model for the organization of the 2G12 dimer that involves intermolecular 3D domain swapping (Fig. 1C). We used the structural model to design and express 2G12 mutants predicted to have a higher fraction of the more potent dimeric form. One of these mutants, D2, had a significantly higher dimer content and equivalent overall expression. Due to the much higher neutralization potencies of the original and the D2 mutant dimers, a modest fraction of dimer can dominate the overall potency of a 2G12 preparation with a mixture of forms. Different 2G12 dimer fractions may occur in different expression contexts, perhaps relating to the density of assembling 2G12 in the endoplasmic reticulum, which could influence the rates of intra- versus intermolecular domain swapping. Further experiments will be required to investigate the extent to which the D2 mutant increases dimer fraction in other expression systems. Although the mechanism by which the D2 mutant causes a higher dimer content is not known for certain, this mutant would be a potentially useful therapeutic reagent in any context in which unfractionated 2G12 is utilized. This includes gene therapy applications, topical (microbicide) use, and use as an injectable reagent.

REFERENCES

1. Bennett, M. J., and D. Eisenberg. 2004. The evolving role of 3D domain swapping in proteins. Structure **12:**1339–1341.
2. Burton, D. R., R. L. Stanfield, and I. A. Wilson. 2005. Antibody versus HIV in a clash of evolutionary titans. Proc. Natl. Acad. Sci. USA **102:**14943–14948.
3. Calarese, D. A., H. K. Lee, C. Y. Huang, M. D. Best, R. D. Astronomo, R. L. Stanfield, H. Katinger, D. R. Burton, C. H. Wong, and I. A. Wilson. 2005. Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. Proc. Natl. Acad. Sci. USA **102:**13372–13377.
4. Calarese, D. A., C. N. Scanlan, M. B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M. R. Wormald, R. L. Stanfield, K. H. Roux, J. W. Kelly, P. M. Rudd, R. A. Dwek, H. Katinger, D. R. Burton, and I. A. Wilson. 2003. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science **300:**2065–2071.
5. Durocher, Y., S. Perret, and A. Kamen. 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. Nucleic Acids Res. **30:**E9.
6. Harris, L. J., S. B. Larson, K. W. Hasel, and A. McPherson. 1997. Refined structure of an intact IgG2a monoclonal antibody. Biochemistry **36:**1581–1597.
7. Harris, L. J., E. Skaletsky, and A. McPherson. 1998. Crystallographic structure of an intact IgG1 monoclonal antibody. J. Mol. Biol. **275:**861–872.
8. Huber, A. H., R. F. Kelley, L. N. Gastinel, and P. J. Bjorkman. 1993. Crystallization and stoichiometry of binding of a complex between a rat intestinal Fc receptor and Fc. J. Mol. Biol. **230:**1077–1083.
9. Janeway, C. A., P. Travers, M. Walport, and M. J. Schlomchik. 2005. Immunobiology, 6th ed. Garland Science Publishing, New York, NY.
10. Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A **47**(Pt. 2):110–119.
11. Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. Sequences of proteins of immunological interest. Department of Health and Human Services, Washington, DC.
12. Li, M., F. Gao, J. R. Mascola, L. Stamatatos, V. R. Polonis, M. Koutsoukos, G. Voss, P. Goepfert, P. Gilbert, K. M. Greene, M. Bilska, D. L. Kothe, J. F. Salazar-Gonzalez, X. Wei, J. M. Decker, B. H. Hahn, and D. C. Montefiori. 2005. Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. J. Virol. **79:**10108–10125.
13. Mehandru, S., B. Vcelar, T. Wrin, G. Stiegler, B. Joos, H. Mohri, D. Boden, J. Galovich, K. Tenner-Racz, P. Racz, M. Carrington, C. Petropoulos, H. Katinger, and M. Markowitz. 2007. Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection. J. Virol. **81:**11016–11031.
14. Montefiori, D. C. 2005. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. Curr. Protoc. Immunol. **Chapter 12:**Unit 12.11.
15. Ramessar, K., T. Rademacher, M. Sack, J. Stadlmann, D. Platis, G. Stiegler, N. Labrou, F. Altmann, J. Ma, E. Stoger, T. Capell, and P. Christou. 2008. Cost-effective production of a vaginal protein microbicide to prevent HIV transmission. Proc. Natl. Acad. Sci. USA **105:**3727–3732.
16. Roux, K. H., P. Zhu, M. Seavy, H. Katinger, R. Kunert, and V. Seamon. 2004. Electron microscopic and immunochemical analysis of the broadly neutralizing HIV-1-specific, anti-carbohydrate antibody, 2G12. Mol. Immunol. **41:**1001–1011.
17. Saphire, E. O., P. W. Parren, R. Pantophlet, M. B. Zwick, G. M. Morris, P. M. Rudd, R. A. Dwek, R. L. Stanfield, D. R. Burton, and I. A. Wilson. 2001. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. Science **293:**1155–1159.
18. Scanlan, C. N., R. Pantophlet, M. R. Wormald, E. Ollmann Saphire, R. Stanfield, I. A. Wilson, H. Katinger, R. A. Dwek, P. M. Rudd, and D. R. Burton. 2002. The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of α1→2 mannose residues on the outer face of gp120. J. Virol. **76:**7306–7321.
19. Sheskin, D. 2004. Handbook of parametric and nonparametric statistical procedures, 3rd ed. Chapman & Hall/CRC, Boca Raton, FL.
20. Stiegler, G., C. Armbruster, B. Vcelar, H. Stoiber, R. Kunert, N. L. Michael, L. L. Jagodzinski, C. Ammann, W. Jager, J. Jacobson, N. Vetter, and H. Katinger. 2002. Antiviral activity of the neutralizing antibodies 2F5 and 2G12 in asymptomatic HIV-1-infected humans: a phase I evaluation. AIDS **16:**2019–2025.
21. Trkola, A., H. Kuster, P. Rusert, B. Joos, M. Fischer, C. Leemann, A. Manrique, M. Huber, M. Rehr, A. Oxenius, R. Weber, G. Stiegler, B. Vcelar,

H. Katinger, L. Aceto, and H. F. Gunthard. 2005. Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. Nat. Med. **11:**615–622.

22. **Trkola, A., M. Purtscher, T. Muster, C. Ballaun, A. Buchacher, N. Sullivan, K. Srinivasan, J. Sodroski, J. P. Moore, and H. Katinger.** 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.

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

24. **Wolbank, S., R. Kunert, G. Stiegler, and H. Katinger.** 2003. Characterization of human class-switched polymeric (immunoglobulin M [IgM] and IgA) anti-human immunodeficiency virus type 1 antibodies 2F5 and 2G12. J. Virol. **77:**4095–4103.