

Structure–Activity Relationships in the Binding of Chemically Derivatized CD4 to gp120 from Human Immunodeficiency Virus

Hui Xie,^{*, ∇} Danny Ng,^{§, \parallel} Sergey N. Savinov,^{§, \perp} Barna Dey,[‡] Peter D. Kwong,[‡] Richard Wyatt,[‡] Amos B. Smith, III,[§] and Wayne A. Hendrickson^{*, \ddagger}

Department of Pharmacology and Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815-6789, and Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032

Received May 15, 2007

The first step in HIV infection is the binding of the envelope glycoprotein gp120 to the host cell receptor CD4. An interfacial “Phe43 cavity” in gp120, adjacent to residue Phe43 of gp120-bound CD4, has been suggested as a potential target for therapeutic intervention. We designed a CD4 mutant (D1D2F43C) for site-specific coupling of compounds for screening against the cavity. Altogether, 81 cysteine-reactive compounds were designed, synthesized, and tested. Eight derivatives exceeded the affinity of native D1D2 for gp120. Structure–activity relationships (SAR) for derivatized CD4 binding to gp120 revealed significant plasticity of the Phe43 cavity and a narrow entrance. The primary contacts for compound recognition inside the cavity were found to be van der Waals interactions, whereas hydrophilic interactions were detected in the entrance. This first SAR on ligand binding to an interior cavity of gp120 may provide a starting point for structure-based assembly of small molecules targeting gp120–CD4 interaction.

Introduction

Human immunodeficiency virus 1 (HIV-1^a) is the primary cause of acquired immunodeficiency syndrome (AIDS).^{1,2} Among 20 antiretroviral drugs that have been approved by FDA for clinical treatment of AIDS,³ nearly all target either the reverse transcriptase or the protease of HIV-1. Currently, the continuing emergence of drug-resistant viruses and drug toxicity problems, as well as the absence of an effective AIDS vaccine, motivates the search for novel antiretroviral drugs.

Elements of virus entry are attractive antiviral targets because they are at the first step of the HIV-1 life cycle, preceding cellular infection. The entry of the virus is mediated by specific interactions between viral envelope glycoproteins (gp120 and gp41) and host cell surface receptors (CD4 and CCR5/CXCR4). HIV-1 virus first attaches to the host cell surface through recognition between gp120 and CD4.^{4,5} CD4 binding induces

extensive structural rearrangements in gp120, resulting in the exposure of a binding surface for the second host cell chemokine receptor, CCR5 or CXCR4,^{6,7} and exposure of new epitopes on gp41.^{8,9} The ensuing engagement of gp120 with the chemokine receptor triggers further conformational changes, which ultimately lead to the fusion of viral membrane with the host cell membrane.^{10–12}

The interface between gp120 and CD4 is highly conserved among different HIV-1 isolates, as evidenced by multiple X-ray crystal structures of core gp120 proteins derived from different isolates in complex with D1D2 (the first two immunoglobulin-like domains of CD4) and a Fab fragment of antibody, 17B or X5.^{13–15} In these complexes, the D1 domain of CD4 binds in a depression on the core gp120 formed by residues from all three domains of gp120, including the inner domain, the outer domain, and the bridging β -sheet. The binding region in CD4 is centered on residues 40–48 of the D1 domain, a stretch corresponding to the second complementarity-determining region (CDR2) of immunoglobulins. Among these residues, Phe43 alone contributes 23% of the total interactions.¹⁴ At the gp120–CD4 interface in all structures, a deep hydrophobic cavity enclosed by conserved gp120 residues has been identified. The CD4 residue Phe43 sits right at the opening of this cavity, acting as a lid, and is the only cavity-interacting residue in CD4. This cavity is thus named “the Phe43 cavity”^{13,14} and has been suggested as a potential target for therapeutic intervention.^{13,14,16} Biochemical,^{16–20} structural,^{21,22} and thermodynamic²³ studies indicate that the Phe43 cavity-bearing CD4-bound gp120 adopts a conformation more rigidified than that of free gp120, and this conformation of gp120 has been estimated to be only sampled by a small fraction (12–15%) of free gp120.²⁴

Most of the potent gp120–CD4 inhibitors identified to date are proteins or peptides.²⁵ A few gp120-directed small-molecular entry inhibitors, e.g., BMS-378806, BMS-488043,^{26–29} and NBD-556,³⁰ have been identified by high-throughput screening, although their modes of action remain controversial.^{31–33}

* To whom correspondence should be addressed. Address: Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032. Phone: (212) 305-3456. Fax: (212) 305-7379. E-mail: wayne@convex.hhmi.columbia.edu.

^{∇} Department of Pharmacology and Department of Biochemistry and Molecular Biophysics, Columbia University.

^{\perp} Present address: Elusys Therapeutics, 25 Riverside Drive, P.O. Box 102, Pine Brook, NJ 07058.

[§] University of Pennsylvania.

^{\parallel} Present address: Medicinal Chemistry, FibroGen Inc., 225 Gateway Boulevard, South San Francisco, CA 94080.

^{\perp} Present address: Department of Chemistry, Purdue University, West Lafayette, IN 47907.

[‡] National Institute of Allergy and Infectious Diseases, National Institutes of Health.

[†] Howard Hughes Medical Institute and Department of Biochemistry and Molecular Biophysics, Columbia University.

^a Abbreviations: HIV-1, human immunodeficiency virus 1; AIDS, acquired immunodeficiency syndrome; CDR, complementarity-determining region; D1D2, two N-terminal IgG-like domains of CD4; IC₅₀, the half-maximal inhibitory concentration; D1D2F43C-“compound X”, modification product of D1D2F43C by compound X; HTS, high-throughput screening; WT, wild type; FL, full length; DTNB, 5,5'-dithiobis(2-nitrobenzoic acid); NEM, *N*-ethylmaleimide; ELISA, enzyme-linked immunosorbent assay.

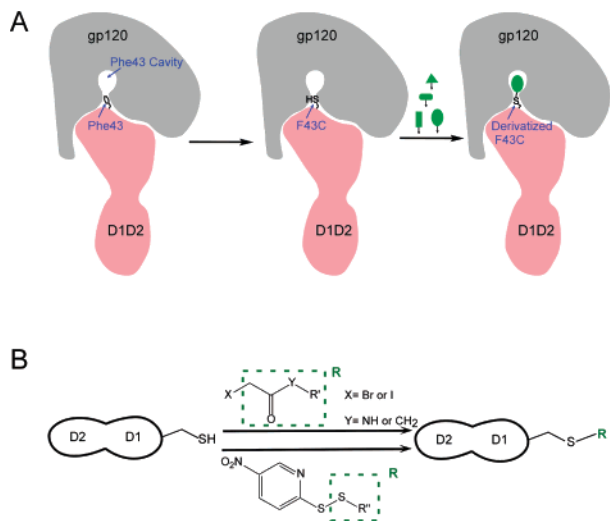


Figure 1. Screen for Phe43 cavity-targeting compounds using derivatized CD4. (A) Design of modified D1D2F43C for targeting the Phe43 cavity in gp120. The complex of gp120 and the D1D2 domains of CD4 are drawn schematically with gp120 and D1D2 colored in gray and salmon, respectively. Residue Phe43, right at the entrance of the Phe43 cavity, is mutated to a chemically reactive cysteine for specific conjugation of a library of cysteine-reactive compounds (in green). The generated D1D2F43C derivatives are then screened for their affinity for gp120. (B) Modification of F43C of D1D2 by haloacetamides or halopropanones (shown on top of the arrows, Y = NH or CH₂, respectively) or 5-nitro-2-pyridinesulfonyl reagents (shown below the arrows).

Possible reasons for the difficulty in identifying a small-molecule inhibiting gp120–CD4 interaction include the unusually large conformational flexibility of gp120 and the intrinsic difficulty of small-molecule drugs to target a complex protein–protein interface,³⁴ such as that between gp120 and CD4. A highly conserved interfacial hydrophobic pocket like the Phe43 cavity in gp120,^{13,14,35} on the other hand, could be a good binding site for small molecules that block protein–protein binding either through a direct steric effect or through an allosteric mechanism. Moreover, such a site should be less prone to resistance-conferring mutations. Thus, in this work, instead of exploring the entire gp120–CD4 interface per se, we characterized the Phe43 cavity as a potential drug-targeting site. Because a large entropic penalty is associated with stabilization of the Phe43 cavity in gp120,²³ direct identification of cavity-targeting small molecules is difficult. By using chemically derivatized CD4 molecules, we could specifically target druglike fragments to the cavity and use them to characterize cavity properties. Herein, we report the construction of 81 modified CD4 molecules and the first structure–activity relationships for the interactions of the Phe43 cavity-targeting probes with gp120.

Results

Design Rationale. The exact location of residue Phe43 of CD4 at the entrance of the cavity makes this locale the best attachment point for cavity-targeting compounds. Cysteine, among all natural amino acids, best enables chemoselective conjugation. Thus, we used a two-domain CD4 (D1D2) scaffold with Phe43 replaced by a cysteine (D1D2F43C) as the template for site-specific delivery of diverse probes to the cavity (Figure 1). D1D2F43C, containing all the essential elements of CD4 in binding gp120 except for the Phe43 side chain, stabilizes gp120 in the cavity-bearing conformation but binds to gp120 much more weakly than CD4 ($IC_{50}(\text{D1D2F43C}) = 206 \pm 21$ nM vs $IC_{50}(\text{D1D2}) = 7.31 \pm 1.07$ nM, Figure 2). This substitution

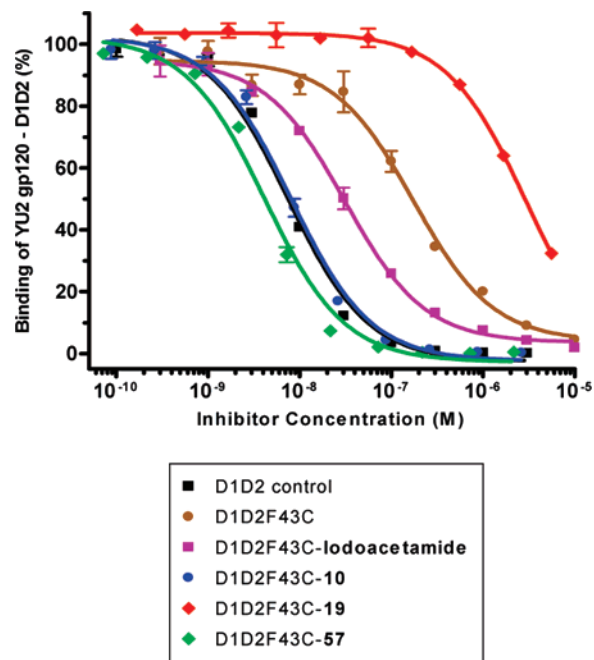


Figure 2. Representative curves for the inhibition of gp120–CD4 binding by D1D2 control, D1D2F43C, and D1D2F43C derivatives (D1D2F43C-iodoacetamide, D1D2F43C-10, D1D2F43C-19, and D1D2F43C-57) in a competition ELISA assay. Curves were generated from binding data using Prism 4 (GraphPad Software). The derived IC_{50} values were 7.31 ± 1.07 , 206 ± 21 , 33.3 ± 5.5 , 7.76 ± 0.83 , 3340 ± 510 , and 4.14 ± 0.03 nM, respectively.

allows the compatibility of a chemical group and the cavity to be easily judged by comparing the gp120-binding affinities of the derivatized D1D2F43C proteins to that of D1D2F43C (Figure 1).

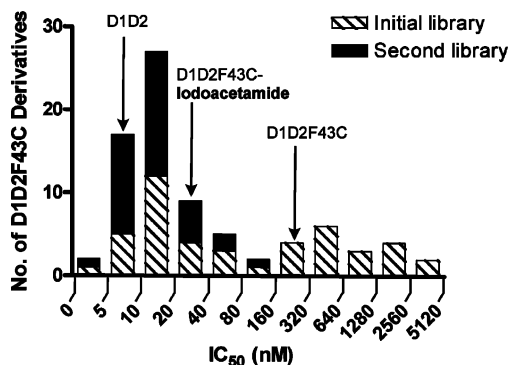
We first tested each of three commonly used classes of sulfhydryl-specific reagents [i.e., iodoacetamide, 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB), and *N*-ethylmaleimide (NEM)] to identify the best thiol-reacting module for linking compounds to Cys43. Although all three compounds were found to modify D1D2F43C with satisfactory completeness (>98%, 80%, and >98%, respectively) (Supporting Information), iodoacetamide was the only reagent that produced a CD4 derivative with gp120-binding affinity higher than that of the unmodified template D1D2F43C, demonstrated by a 6-fold lower IC_{50} value (Figure 2 and Supporting Information). Consequently, haloacetamide was chosen as the primary starting module for construction of the first thiol-reactive compound library, which included 41 bromoacetamides (compounds **1** (SNS-1) to **41** (SNS-41)) and 7 5-nitro-2-pyridyldisulfides (compounds **42** (SNS-42) to **48** (SNS-48)), designed through a computer-assisted molecular complementarity search using GrowMol.^{36,37} On the basis of the screening results from the first library, we further designed a second library of 42 bromoacetamides, 1 bromopropanone, 1 bromoethylamine, and 2 5-nitro-2-pyridyldisulfides (compounds **49** (DN-10) to **94** (DN-271)) to complete the analysis of structure–activity relationships of cavity-filling probes (Table 1 in Supporting Information).

All D1D2F43C derivatives were generated employing nucleophilic reactions between thiolate anions of D1D2F43C and electrophiles of the two compound libraries synthesized (Figure 1). The derivatized proteins were named D1D2F43C-“compound” according to the compounds used to derive them; e.g., D1D2F43C-1 is the product of the modification of D1D2F43C by compound **1**. A “D1D2 control” was prepared as a modifica-

Table 1. Summary of Modification of D1D2F43C by Both Compound Libraries^a

cysteine-modifying compound	initial library		second library		total	
	reacted, no.	unreacted, no.	reacted, no.	unreacted, no.	reacted, no.	unreacted, no.
haloacetamides	39	2	35	9	74	11
disulfides	6	1	1	1	7	2
total	45	3	36	10	81	13

^a Reacted: completeness, $\geq 80\%$. Unreacted: completeness, $< 80\%$.

**Figure 3.** Distribution of the IC_{50} values of the D1D2F43C derivatives generated from the initial and second libraries.

tion control by using compound **40** to “mock-modify” wild type (WT) D1D2 in the same way as described above.

Overview of Binding Activities of CD4 Derivatives. Out of 94 compounds synthesized, 81 modified D1D2F43C with $> 80\%$ completeness (Table 1 and Supporting Information). The resulting 81 derivatives were tested for their abilities to inhibit the interaction of gp120 with D1D2 in a competition ELISA assay (Figure 2), from which an IC_{50} value (the concentration for obtaining 50% inhibition of gp120–CD4 binding) was determined for each of the derivatives. Because the D1D2F43C derivatives bind gp120 competitively with D1D2 and the IC_{50} value of the D1D2 control is in good agreement with the reported K_d of D1D2 and gp120,^{38,39} the measured IC_{50} values were regarded as the estimations of affinities between the derivatives and gp120.

The derivatives from the initial library had widely spread IC_{50} values, whereas those from the second library, optimized from results of initial libraries, all bound gp120 better than D1D2F43C (Figure 3). Furthermore, 90% of the proteins derivatized from the second library had IC_{50} values lower than that of D1D2F43C-iodoacetamide, the parent derivative for all haloacetamide-derived proteins. Out of all derivatives, eight had IC_{50} values lower than that of D1D2 (7.3 nM). The best derivative was D1D2F43C-**57**, with an IC_{50} value almost half that of D1D2 (Figure 2). The affinity of the derivatized D1D2F43C proteins for gp120, however, reached a plateau with an IC_{50} value of around 4 nM even after extensive optimization.

Validation of the Binding Sites for CD4 Derivatives. We confirmed that the chemical probes presented by CD4 derivatives bind into the Phe43 cavity of gp120 by comparing the affinities of the derivatives for gp120 to their affinities for a Phe43 cavity-filled mutant of gp120. The cavity-filled gp120 mutant, namely, gp120 S375W/T257S, has the cavity-lining residue Ser375 mutated to a tryptophan, which fills the cavity (Figure 4A).^{20,22,50} The other mutation, T257S, was incorporated to stabilize gp120 S375W. Unmodified D1D2 and D1D2F43C, which do not bind into the cavity, bound the cavity-filled mutant of gp120 equally well as WT gp120, as indicated by the identical IC_{50} values in competition assays (Figure 4B,C). In contrast, all except one of the D1D2F43C derivatives, regardless of their affinities for gp120, exhibited dramatically lower affinities (4-

to 150-fold higher IC_{50} values) for gp120 S375W/T257S than for WT gp120 (Figure 4B,C), which is consistent with the Phe43 cavity being the binding sites for the chemical groups linked to Cys43 in these derivatives. Among the derivatives with nanomolar IC_{50} values, the molecular weights of the compounds were found to correlate well with the loss of affinity for cavity-filled gp120 (data not shown). D1D2F43C-iodoacetamide, the only derivative with equal affinity for gp120 S375W/T257S and WT gp120, had only a small acetamide moiety attached to residue Cys43, which was not expected to protrude into the cavity based on our modeling results (data not shown).

SAR of Binding of CD4-Attached Compounds to gp120: N-Alkylacetamide Derived D1D2F43C. Nine derivatives were evaluated for their gp120-binding affinities to study the effect of the addition of alkyl groups at the acetamide nitrogen of the parent derivative D1D2F43C-iodoacetamide (Table 2). Small alkyl groups such as isopropyl (**1**) reduced the affinity greatly (5-fold), whereas larger alkyl groups such as isobutyl (**2**) and 6-hydroxyhexyl (**9**) did not have much affect on the affinity. An even bulkier cyclic group [i.e., the cyclohexyl group (**3**)] increased the binding affinity more than 2-fold with an IC_{50} of 13 nM. This positive effect was further enhanced when an extra methylene linker was inserted between the acetamide nitrogen and the cyclohexyl group to yield the cyclohexylmethyl group (**4**). Addition of even bigger groups (**6**, **56**, and **8**) reduced the binding affinities drastically to micromolar levels, suggesting that the volume limit of the cavity had been reached.

SAR: N-Arylacetamide Derived D1D2F43C. Unlike alkyl groups, aryl groups were much more favored when attached to the acetamide nitrogen. The highest binding affinity in this group (Table 3) was observed with D1D2F43C-**10** ($IC_{50} = 7.76$ nM), which has a phenyl group at the acetamide nitrogen. Insertion of up to three methylene linkers between the nitrogen and phenyl groups (**28**, **40**, and **51**) decreased binding affinity slightly with an ethylene group being the best linker. Interestingly, branching at the benzylic position resulted in a drastic (5- to 10-fold) loss in affinity (**29** and **30** compared to **28**). Similar to the phenyl group, a naphthyl group (**14**) at the acetamide nitrogen was also favorable, but the insertion of additional linkers (**32**, **64**, **31**, and **66**) between the naphthyl unit and the acetamide nitrogen attachment was much less tolerated.

In light of the preference for phenyl- and naphthyl-like groups in the cavity, we also synthesized compounds containing five-membered aromatic rings linked to the acetamide. The findings (Table 4) were very similar except for compound **67**, the derivative of which has affinity 3-fold lower than those of the derivatives of **10** and **87**, suggesting possibly unfavorable interaction between gp120 and the sulfur atom at position 2 in thiophen.

SAR: Substituted N-Phenylacetamide Derived D1D2F43C. Extensive substitutions on the phenyl group of compound **10** were employed both to screen for higher-affinity binders compared to the parent compound **10** and to characterize further the chemical preferences of the Phe43 cavity. Eighteen compounds were synthesized with various substitutions at the para position (Table 5A). In general, small substitutions ranging from

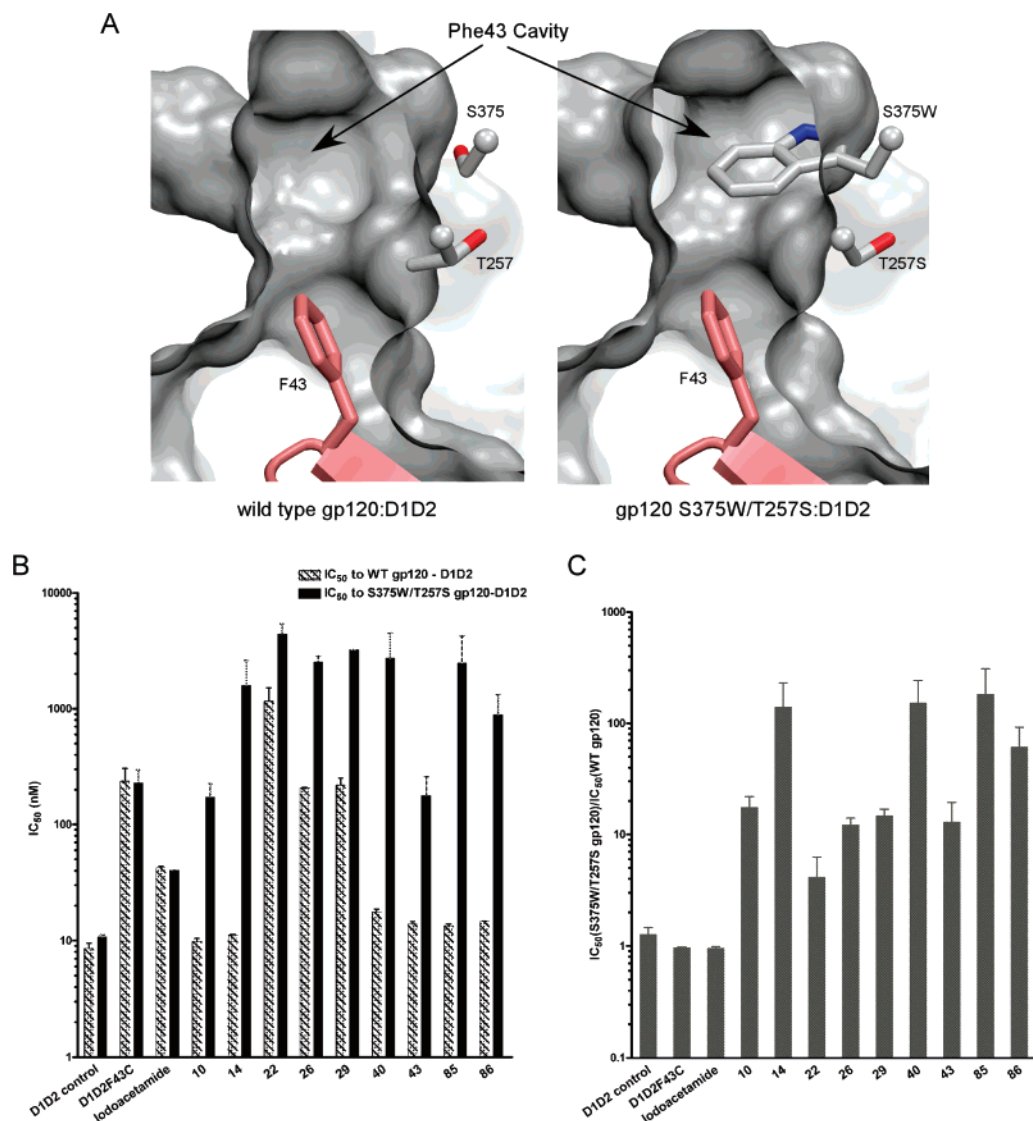


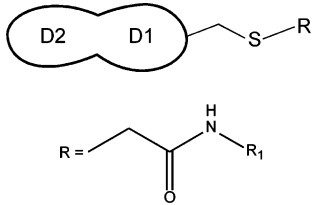
Figure 4. Binding of D1D2 conjugates to cavity-filled gp120. (A) Sliced-open surface representations of the Phe43 cavities of wild type core gp120 (left, PDB code 1RZJ) and cavity-filled core gp120 mutant S375W/T257S (right, PDB code 2NXZ) bound to CD4. CD4 and gp120 molecules are colored in gray and salmon, respectively. Side chains and C α atoms of gp120 residues 375 and 257, as well as those of CD4 residue F43, are shown in ball-and-stick model in the coloring scheme: CD4 carbon atoms (salmon), gp120 carbon atoms (gray), nitrogen atom (blue) and oxygen atoms (red). Surfaces in both panels were calculated based on the respective gp120 model in which the side chains of residues 375 and 257 except for C β atom were removed. Note that the cavity has a narrow entrance close to the tip of phenyl ring of F43 and that a water channel is located left to the cavity. (B) Side by side comparisons of IC₅₀ values of D1D2 variants on the binding of D1D2 to YU2 WT (wild type) gp120 to their IC₅₀ values on the binding of D1D2 to YU2 S375W/T257S gp120. (C) Ratios of two IC₅₀ values.










methyl (**73**) to ethoxy (**74**) groups had limited effect on the activities, but larger groups such as isopropoxy (**22**) and benzyloxy (**26**) groups reduced the affinities significantly. The highest binding affinities of this whole study were observed when compound **10** was substituted at the para position by a nitro (**57**, 4.14 nM), an isopropyl (**12**, 4.84 nM), or an ethyl (**76**, 5.87 nM) group. Interestingly, derivatives with hydroxy-containing substituents at the para position invariably showed IC₅₀ values higher than those of similar derivatives with the hydroxy groups replaced by the methyl groups (**59** and **73**; **77** and **76**). This finding indicates a disfavoring of electron-donating groups at the para position. An isopropanol has been reported to occupy the Phe43 cavity in the refined 2.2 Å structure of HXBc2 gp120:D1D2:17b Fab,¹³ suggesting a possible starting point for the design of cavity-binding compounds. Substitution of **10** at the para position by an isopropanol group (**94**) led to slightly reduced affinity, indicating disfavor of the isopropanol group at this orientation.

Eight groups were tested for meta-position substitution of the phenyl ring in compound **10**. Five of these (methyl, hydroxy, nitro, 2-hydroxyethoxy, and benzyloxy groups) were also tested at the para and ortho positions. A much higher tolerance for bulkier groups was noticed at the meta position compared to the para position (Table 5B). Smaller groups at the meta position, on the other hand, bound comparably well or somewhat worse (nitro) than at the para position.

None of seven different substitutions at the ortho position of the phenyl group was able to enhance the affinity (Table 5C). Interestingly, the smallest group tested, methyl (**78**), was the second to the least favored at this position with an IC₅₀ of 24.5 nM; yet a much bulkier group benzyloxy (**85**) performed best at this position with an IC₅₀ of 8.87 nM, comparable to that of the parent compound **10**.

SAR: Role of the Acetamide Moiety. Modification of D1D2F43C by the simplest haloacetamide-based compound that we tested (iodoacetamide) increased the binding affinity for

Table 2. Chemical Structures of Cysteine-Modifying Compounds (R as *N*-Alkylacetamide) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a


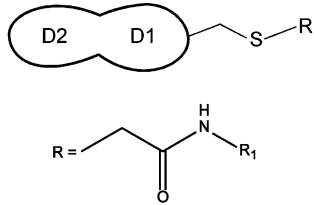
Compound	R ₁	IC ₅₀ ± SD of Derivatized D1D2 (nM)
Iodoacetamide	H	33.3 ± 5.5
1		184 ± 39
2		29.4 ± 2.2
9		33.0 ± 9.7
3		13.0 ± 2.1
4		7.79 ± 2.08
5		17.0 ± 2.5
6		223 ± 49
56		42.3 ± 5.9
8		1680 ± 500












^a Corrected IC₅₀ values are presented as the mean ± SD values from two to three independent experiments.

gp120 by 6-fold (Figure 2). On the basis of our modeling result (data not shown), the acetamide moiety binds at the vestibule to the Phe43 cavity, which is the binding site for residue Phe43 of WT CD4. The 6-fold affinity increase suggests a good complementarity of the acetamide group to the vestibule, possibly due to the relatively small size as well as rigid and nonbranching shape. The narrower shape of the acetamide moiety compared to the phenyl ring of Phe43, however, probably limits the ability to make extensive interactions with gp120 around the cavity opening, in contrast to Phe43, as suggested by affinity of D1D2 that is 5-fold higher than that of D1D2F43C-iodoacetamide.

We used compound **10** as a positive background to identify the gp120-binding elements of the acetamide group. Replacement of the acetamide nitrogen by a carbon (**71**) decreased the binding affinity by 3-fold (Table 6), indicating that the acetamide nitrogen plays a significant role in the interaction of the acetamide moiety with gp120, probably through an interaction for which the methylene carbon cannot substitute, such as the hydrogen bonding. Compound **69**, in which the carbonyl oxygen of the acetamide moiety was removed, failed to modify D1D2F43C (Supporting Information). Therefore, the significance of the carbonyl oxygen in the acetamide is unclear. Some indication may be derived from comparison of **71** with a disulfide compound **63**, which resembles **71** without the carbonyl oxygen (Table 6). The similar IC₅₀ values of these two derivatives suggest either that the carbonyl oxygen is not important for the interaction or that its deletion is compensated in some way by substitution of sulfur for carbon.

SAR: Mixed-Disulfide D1D2F43C Derivatives. Only a small series of disulfide compounds were investigated in this study because of their limited solubility and poor reactivity. All

Table 3. Chemical Structures of Cysteine-Modifying Compounds (R as *N*-Arylacetamide) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a


Compound	R ₁	IC ₅₀ ± SD of Derivatized D1D2 (nM)
10		7.76 ± 0.83
28		15.2 ± 0.7
29		88.7 ± 1.5
30		166 ± 22
40		10.6 ± 2.9
51		16.4 ± 2.6
14		10.4 ± 0.8
32		75.7 ± 11.2
64		69.3 ± 1.8
31		12.2 ± 0.8
66		146 ± 4

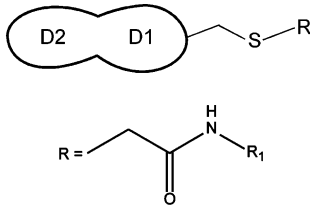
^a Corrected IC₅₀ values are presented as the mean ± SD values from two to three independent experiments.






derivatives contain aryl groups that were linked to Cys43 through disulfide bonds (Table 7). The incorporation of most of these aryl groups (6 out of 7) significantly increased activities of derivatives in binding gp120 when compared with D1D2F43C. Similar to the SAR in the substituted *N*-phenylacetamide derived D1D2F43C, addition of an isopropyl group but not other bulkier groups at the para position of the benzyl group in compound **42** slightly enhanced affinity (**43**).

Discussion

Screen for Phe43 Cavity-Targeting Inhibitors Using Derivatized CD4. The accessible surface of gp120 is thought to be largely composed of both highly variable and heavily glycosylated regions,^{14,16} and the conserved Phe43 cavity is thought to be poorly formed in free gp120.²⁴ Hits identified in high-throughput screens against gp120 are likely to recognize variable regions, and thus, the resulting compounds are expected to be highly susceptible to resistance because of high mutation rates. In contrast, the emergence of resistant viruses to inhibitors targeting the conserved Phe43 cavity should be rare, if at all possible. The Phe43 cavity is thought to exist preferably in the relatively rigidified gp120 state when bound to CD4, and it is disrupted in the free SIV core gp120²¹ and the HIV gp120 complex with Fab b12.²² Mutation of the cavity-lining residue Ser375 to a bulky tryptophan fills the Phe43 cavity and drives gp120 into a conformation closer to that of the CD4-bound state.²⁰ The large entropy change associated with the transition

Table 4. Chemical Structures of Cysteine-Modifying Compounds (R as *N*-Thiophen(furan)acetamide) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a



Compound	R ₁	IC ₅₀ ± SD of Derivatized D1D2 (nM)
87		7.40 ± 1.04
67		18.2 ± 1.8
68		13.1 ± 0.5
53		18.3 ± 2.1
55		18.8 ± 0.5

^a Corrected IC₅₀ values are presented as the mean ± SD values from two to three independent experiments.

of free gp120 to CD4-bound gp120²³ undoubtedly determines that gp120 poorly populates cavity-bearing conformations,²⁴ which reduces the chance of identifying hits targeting the Phe43 cavity in conventional high-throughput screening (HTS).

New methods for identifying lead compounds or fragments of lead compounds as inhibitors of protein–protein interactions have emerged recently as alternatives to HTS. These include structure-based screening by NMR techniques and X-ray crystallography, in silico screening, phage display, rational discovery, and tethering.^{40,41} One of these approaches, tethering,⁴² identifies small-molecule leads by screening the compounds covalently linked to the targeting site. This method, however, is unlikely to be applicable for identifying inhibitors for gp120 because of the high intrinsic flexibility of this target. Moreover, highly glycosylated gp120 does not readily permit the use of mass spectrometry in identifying suitable compounds in the tethering method. Unlike other approaches, our method focuses on the screening of the compounds that are covalently attached to a protein ligand (CD4) against the target protein (gp120). CD4, as the natural ligand for gp120, stabilizes formation of the Phe43 cavity in gp120. In addition, using a gp120 ligand such as CD4 as the ligand-attachment point in our method enables the quantitative determination of structure–activity relationships on the small-molecule fragments, which would otherwise bind gp120 too weakly for accurate affinity determination using traditional binding assays. A parallel study recently proposed a similar idea; the protocol entailed assembly of a library of peptide–small molecule hybrids via cysteine modification for studying SAR between the target protein and small molecules;⁴³ however, to date, no screening results have been reported.

Although our results are fully consistent with binding small molecules into the specific target site, surely the CD4 scaffold to which these chemicals are attached must also contribute to the binding properties. Scaffold constraints also affect binding properties in more conventional combinatorial chemistry approaches, but in our approach constraints from the larger D1D2 moiety might have a more substantial impact. Various aspects

of the data nevertheless indicate that the SAR is indeed reflective of intrinsic molecular complementarity between the compounds and the cavity, at least for those derivatives with higher affinities than the unmodified template, D1D2F43C. The observed binding affinities for gp120 are widely distributed for the various chemically derivatized CD4 molecules, and these affinities are drastically lowered for cavity-filled gp120. This proves that the binding affinities of the derivatives to gp120 are sensitive to the interaction between the template-attached cavity-probing compounds and the cavity. Moreover, in the one possibility for comparison, this sensitivity is independent of the scaffold. That is, the enhancement (4-fold) of gp120-binding affinity upon addition of a cavity-targeting phenyl group to the vestibule-binding template D1D2F43C-iodoacetamide (vs D1D2F43C-10) is found to be the same as that gained (4-fold) through addition of a phenyl group to the existing phenyl group equivalent of that of Phe43 in a CD4 miniprotein mimetics F23.³⁵ D1D2F43C and the miniprotein scaffold are radically different in size, have different vestibule-binding groups, and have substantially different intrinsic affinities; thus, the high similarity in relative binding characteristics suggests that high-affinity derivatives probably bind the cavity with minimum constraints from the template. Therefore, the SAR derived from this study likely reflects true binding preferences of the Phe43 cavity.

Characterization of the Phe43 Cavity. The Phe43 cavity has been observed in the crystal structures of CD4-bound core gp120^{13–15} to have an entrance narrower than the main body of the cavity (Figure 4A). Modeling of the acetamide group onto the Cys43 of D1D2 positioned the acetamide nitrogen close to the narrow cavity entrance of core gp120 (data not shown). Disfavoring of branching at the first carbon linked to the acetamide nitrogen, such as seen with compounds **1**, **29**, and **30** (Tables 2 and 3), suggests that the narrow cavity entrance is also present in the full-length (FL) gp120 used in this SAR study and that the acetamide nitrogen atom is near this constriction. Hydrophilic interaction, such as a hydrogen bond, is likely to be part of the recognition mechanism at this site, as discussed earlier. In the disulfide-linked derivatives, branching at the carbon next to the sulfur also had a negative effect on binding affinity (**44**, Table 6). Although the branching point is further away from the cavity than the acetamide nitrogen, the large branched phenyl group in **44** could potentially reach the cavity entrance and cause unfavorable steric clashes. The repulsions caused by the branching at the cavity entrance could be partially released by insertion of a methylene linker between the acetamide nitrogen and the branching groups, as demonstrated by compounds **6** and **56** (Table 2). Preference of aromatic groups over aliphatic groups by the cavity (Tables 2–7) is also in agreement with the presence of a narrow entrance to cavity where the rigid and planar shapes of aromatic groups should be preferred. Additional methylene linkers between the nitrogen atom and aromatic groups, unlike in the case of the alkyl groups, affected binding adversely (Tables 3 and 4), albeit only slightly. One plausible explanation is that while the phenyl group fits well at the narrow cavity entrance, the aromatic ring can no longer maintain all the favorable van der Waals contacts once positioned in the broader space inside the cavity because of the insertion of methylene linkers.

The para position was found to be the best position to substitute the aromatic groups, such as the phenyl ring in compound **10** (Table 5), for increase of gp120-binding affinity. The best substituents found are nitro (**57**) and isopropyl (**12**) groups. Although the groups possess different electronic proper-

Table 5. Chemical Structures of Cysteine-Modifying Compounds (R as Substituted *N*-Phenylacetamide) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a

A. para substituents			B. meta substituents		
Compound	R ₂	IC ₅₀ ± SD of Derivatized D1D2 (nM)	Compound	R ₂	IC ₅₀ ± SD of Derivatized D1D2 (nM)
73	—	7.07 ± 1.51	81	—	6.59 ± 0.89
59	—OH	11.4 ± 0.2	82	—OH	9.82 ± 1.92
50	—F	7.08 ± 1.24	79	—NO ₂	7.12 ± 1.3
76		5.87 ± 0.17	93		11.4 ± 0.1
77		10.6 ± 3.0	89		14.9 ± 0.4
12		4.84 ± 0.83	92		22.0 ± 0.4
13		8.14 ± 0.38	84		14.3 ± 0.25
57	—NO ₂	4.14 ± 0.03	91		36.2 ± 6.5
54		9.87 ± 0.67	C. ortho substituents		
94		10.9 ± 1.3	Compound	R ₂	IC ₅₀ ± SD of Derivatized D1D2 (nM)
74		7.23 ± 0.14	78	—	24.5 ± 1.0
22		1500 ± 330	83	—OH	9.06 ± 0.41
88		16.4 ± 0.7	80	—NO ₂	17.8 ± 4.0
27		1210 ± 340	86		12.9 ± 1.4
23		637 ± 202	90		17.6 ± 1.7
17		402 ± 68	15		51.9 ± 6.3
16		46.2 ± 1.3	85		8.87 ± 1.09
26		290 ± 84			

^a Corrected IC₅₀ values are presented as the mean ± SD values from two to three independent experiments.

ties but similarities in molecular shape, the similar effects on binding of these two groups suggest that they may enhance binding solely through van der Waals interaction with cavity residues. This finding, together with the adverse effect on the affinity of many hydrophilic substituents on the aromatic ring (Tables 4 and 5), indicates that van der Waals interactions may dominate the recognition of the tethered compounds within the cavity. This conclusion is consistent with the hydrophobic nature of the cavity seen in the crystal structures of core gp120.^{13,14} The finding that an isopropanol molecule occupies the cavity in the crystal structure of core gp120¹³ is not inconsistent with this observed preference for hydrophobic groups in the cavity because the hydroxy group of isopropanol is hydrogen-bonded to two water molecules but not to any residue in gp120.

Although substitutions at the meta and ortho positions of the phenyl ring in compound **10** did not result in the best binding to gp120, much larger groups than those allowed at the para position can be substituted at these positions without affecting binding affinity (Table 5). This finding suggests that the meta- or


Table 6. Chemical Structures of Cysteine-Modifying Compounds (Role of the Acetamide Moiety) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a








Compound	R	IC ₅₀ ± SD of Derivatized D1D2 (nM)
10		7.76 ± 0.83
63		24.7 ± 8.6
71		26.4 ± 3.9

^a Corrected IC₅₀ values are presented as the mean ± SD values from two to three independent experiments.

ortho-substituted phenyl groups may bind in the cavity in differ-

Table 7. Chemical Structures of Cysteine-Modifying Compounds (Mixed-Disulfide Compounds) and IC₅₀ Values of the Corresponding Derivatized D1D2F43C for Inhibition of the Binding of D1D2 to YU2 gp120^a



Compound	R	IC ₅₀ ± SD of Derivatized D1D2 (nM)
42		11.5 ± 1.8
43		9.14 ± 1.99
44		1380 ± 100
45		14.3 ± 0.6
47		27.0 ± 1.6
48		14.9 ± 0.0
63		24.7 ± 8.6

^a Estimated IC₅₀ values are presented as the mean ± SD values from two to three independent experiments. Correction for compound absorbance was not possible here because of absorbance from the 5-nitro-2-pyridyl leaving group.

ent modes from the para-substituted phenyl groups. In addition, the binding may involve expansion of the cavity region in gp120.

As discussed, the Phe43 cavity is thought to be conserved in gp120 because it is composed of highly conserved residues across HIV-1.^{13,14} However, whether the structure of the cavity in core gp120 resembles that in the FL gp120 was not clear. Our SAR study of the binding of the probes to the Phe43 cavity in FL gp120 now reveals a close similarity (narrow entrance and hydrophobic surface) between the cavities in FL and core gp120 and thus suggests that the core gp120 structure holds great promise for development of structure-based inhibitors targeting FL gp120.

In addition, our SAR analysis implies significant plasticity in gp120. All derivatives except for D1D2F43C-iodoacetamide were found to bind inside the Phe43 cavity, despite the fact that the derivatization groups on some of the derivatives are much larger than the size of the cavity seen in the structures of core gp120.^{13–15} Whether all parts of each compound could fit completely in the cavity is not clear, but it is reasonable to believe that gp120, even when bound to CD4, has substantial plasticity to undergo necessary structural rearrangements around the Phe43 cavity to accommodate bulky compounds, such as **85** (Table 5C), a derivative that maintains high affinity for gp120 (IC₅₀ = 9 nM).

Future Development and Therapeutic Applications. Although gp120 retains significant plasticity during ligand binding in the presence of the CD4 scaffold, gp120 adaptability is nevertheless restricted; thus, potentially favorable interactions of compounds bound into the intrinsically flexible Phe43 cavity may not be realized fully. This proposal may explain the IC₅₀ plateau of 4 nM reached after optimization of the derivatives. The potential penalty from occupying a cavity that is intrinsically flexible but stabilized by the D1D2 scaffold may prevent a further increase in affinity of derivatized D1D2F43C. A smaller CD4 scaffold such as a mimetic may comprise a more

reasonable template for further optimization of cavity-binding fragments because it may structure gp120 to a lesser degree.

When bound to CD4 derivatives, gp120 probably adopts a rigidified conformation close to that of the CD4-bound form. CD4 derivatives, however, may not fully activate gp120 into a chemokine-receptor binding conformation because a cavity-binding CD4-mimetic CD4M33 (a scyllatoxin-based miniprotein with a biphenylalanine residue at the position corresponding to Phe43 of CD4) has been shown to promote chemokine-receptor binding 6-fold less than does CD4⁴⁴ and to induce 40% less entropic change in gp120 compared to F23, which has a phenylalanine replacing biphenylalanine in CD4M33.³⁵ Identification of cavity-lining residues critical for the inhibitory effect of BMS-368806, a small molecule that inhibits HIV-1 entry most likely by blocking CD4-induced conformational changes in HIV glycoproteins,³¹ further supports that cavity-targeting CD4 derivatives may serve as entry inhibitors both by preventing viral attachment to CD4 and by partially precluding chemokine-receptor induced viral fusion. The immediate clinical application of the cavity-targeting fragments identified in this study would include incorporation of these moieties into current CD4-based therapeutic agents, such as CD4–IgG₂^{45–47} and CD4–Ig,⁴⁸ for possibly better clinical efficacy.

Conclusion

We have designed and constructed a library of chemically derivatized CD4 and carried out a structure–activity study of these compounds binding to the Phe43 cavity in gp120. The Phe43 cavity in FL gp120 was found to be similar to that observed in the crystal structure of core gp120. It is hydrophobic and plastic in nature and has a narrow entrance, where hydrophilic interactions may play a significant role in binding ligands. Aromatic groups attached to an acetamide moiety and substituted with small groups at the para position were found to be most preferred by the cavity. The detailed characterization of the cavity and identification of the chemical preferences for cavity-binding ligands should benefit further development of the identified compounds into efficient gp120–CD4 inhibitors by incorporating them into either smaller protein-like scaffolds (e.g., miniproteins) or small molecules by fragment assembly. The unique utilization of a reactive ligand template (D1D2F43C) in screening compounds for targeting a receptor site (the Phe43 cavity in gp120) at the ligand–receptor interface may be applied to other systems where the stabilization of the targeting site on the receptor requires the presence of a ligand.

Experimental Section

Thiol-Reactive Reagents. Iodoacetamide, 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB), and *N*-ethylmaleimide (NEM) were purchased from Sigma-Aldrich. All other cysteine-modification compounds were synthesized in one to seven steps using commercially available compounds. Experimental details of the syntheses and characterization of these compounds are included in the Supporting Information.

Protein Reagents. The gp120 antibody 17b was produced in ascites and purified by Strategic BioSolutions (DE). Full-length YU2 gp120 was produced in *Drosophila* S2 cells as described previously⁴⁹ and purified similarly as described before.^{13,14} WT and S375W/T257S YU2 were expressed in serum-free medium by transient transfection of HEK293T cells and were purified by either F105 (WT YU2) or 17b (S375W/T257S YU2) affinity columns as previously described.⁵⁰ The numbering of gp120 amino acid residues is based on the sequence of the prototypic HXBc2 strain.⁵¹

Preparation of Recombinant D1D2 and D1D2 Mutants. Recombinant two-domain CD4 (D1D2, residues 1–183) was cloned into NcoI and XhoI sites of vector pET24d (Novagen). D1D2 was

expressed as inclusion bodies in Rosetta (DE3) cells (Novagen) through leaky expression of T7lac promoter without IPTG induction in SuperBroth medium (BIO 101, Inc.) with addition of 0.2% glucose and 20 mM MgCl₂ at 37 °C for 24 h. The inclusion bodies were isolated from cells by sonication and centrifugation and were then washed three times with 2% Triton-X 100, 2 M urea, 5 mM EDTA, and 5 mM DTT in Tris-HCl buffer, pH 7.5. Purified inclusion bodies were solubilized by 6 M guanidine-HCl, 5 mM EDTA, 20 mM Tris-HCl, pH 7.5, 10 mM DTT, and 0.5 mM PMSF. Solubilized D1D2 proteins were refolded at 0.5 mg/mL in a refolding solution optimized based on the no. 10 condition of Foldit kit (Hampton Research): 50 mM Tris-HCl, pH 8, 10 mM NaCl, 1 mM KCl, 1 mM EDTA, 440 mM sucrose, 2 mM cysteine, and 0.5 mM cystine. Impurities and oligomeric D1D2 were removed by passing refolded D1D2 proteins through Q and SP Sepharose Fast Flow resins (Amersham) in batch mode at pH 10.5 and pH 6.2, respectively. As the last step, a size exclusion column Superdex 200 26/60 (Amersham) was used to separate any residual misfolded oligomeric D1D2 from monomeric D1D2. Purified soluble D1D2 contains residues 1–183 of CD4 and an additional glycine from the cloning vector at the N-terminus. D1D2F43C was created by site-directed mutagenesis and prepared similarly as WT D1D2.

Preparation of Derivatized D1D2F43C. Iodoacetamide, DTNB, and NEM were dissolved and stored as 20 mM solution in 0.5 M Na/K phosphate, pH 7.4. All other compounds including both bromo compounds and mixed-disulfide compounds were dissolved in DMF, DMSO, or ethanol as 20–60 mM stock. D1D2F43C protein at 1–3 mg/mL was first reduced by 2 mM dithiothreitol (DTT). Excessive DTT was then removed by PD-10 columns (Amersham), and at the same time the proteins were exchanged into proper reaction buffer: 0.1 M phosphate buffer, pH 7.4, 0.1 M NaCl, 1 mM EDTA, and 0.1 mM DTT for all bromo compounds, iodoacetamide, and DTNB; 0.1 M phosphate buffer, pH 7.4, 0.1 M NaCl, and 1 mM EDTA for 5-nitro-2-pyridinesulfonyl compounds; and 0.1 M phosphate buffer, pH 7.0, 0.1 M NaCl, 1 mM EDTA, and 0.1 mM DTT for NEM.⁵³ The thiol-reactive compounds were diluted into protein solutions to reach a final reaction concentration of 2 mM, 10 times excess of the free thiol groups in the reaction. The reactions were allowed to continue at 25 °C for 2 h in the dark. The final products, namely, derivatized D1D2F43C proteins, were separated from small molecules by passage through desalting columns (PD-10 columns, Amersham) followed by solvent exchange in Amicon Ultra-4 5K concentrators (Millipore) with reaction buffer 3 times. The completeness of the modifications was examined using protein mass spectrometry (MALDI-TOF, Columbia University Protein Core and DNA Sequencing Facility) as well as peptide mass spectrometry of trypsin-digested proteins (Columbia University Protein Core and DNA Sequencing Facility). This latter step was also used to confirm the correct site-directed incorporation of the compounds.

Concentration Determination of Derivatized D1D2F43C Proteins. The “estimated concentrations” of proteins were calculated from 280 nm absorbance of protein samples in PBS using the theoretic extinction coefficient of D1D2F43C. The concentrations of D1D2F43C variants derived from the halo compounds were then corrected by the corresponding correction factors:

$$\epsilon_{\epsilon, R}^{\text{measured}} = \frac{\epsilon_{\text{D1D2F43C}}^{\text{calcd}}}{\epsilon_{\text{D1D2F43C}}^{\text{calcd}} + \epsilon_{\text{R}}^{\text{measured}}}$$

where $\epsilon_{\text{D1D2F43C}}^{\text{calcd}}$ is the theoretic extinction coefficient of D1D2F43C and $\epsilon_{\text{R}}^{\text{measured}}$ is the experimentally determined extinction coefficient of thiol-reacting compounds in PBS. The “estimated concentration” was used for derivatized D1D2F43C from all other compounds.

The concentrations of a selected group of D1D2F43C derivatives were also determined by amino acid compositional analysis (Keck Biotechnology Resource Laboratory, Yale University) and were found to agree well with the results obtained by the UV absorbance method discussed above.

Competition ELISA Assay. The abilities of derivatized D1D2F43C proteins to bind gp120 and inhibit gp120–CD4 interaction were evaluated using a competition ELISA (enzyme-linked immunosorbent assay). Briefly, Immuno 2HB plates (Thermo LabSystem) were coated with 100 μ L of 4 μ g/mL recombinant D1D2 in coating buffer (10 mM phosphate buffer, pH 7.4, and 0.15 M NaCl) overnight at 4 °C. The plates were then blocked by 3% bovine serum albumin (BSA) (CalBiochem) in PBS (300 μ L per well) for 2 h at 25 °C. Fifty nanograms of YU2 gp120 diluted in 100 μ L of 3% BSA-PBS was added to the plates in the absence or presence of one of the D1D2 variants (either D1D2 or a D1D2F43C derivative) at various ranges of concentrations and incubated for 90 min at 25 °C. After removal of unbound gp120 by a four-time wash with PBST (0.05% Tween-20 in PBS), the bound gp120 was detected by a gp120 antibody 17b (100 μ L, 1 μ g/mL), which was further probed with a peroxidase-conjugated donkey antihuman antibody (Jackson ImmunoResearch, 1:20000 dilution, 100 μ L). 3,3',5,5'-Tetramethylbenzidine (Sigma) was used as the substrate for peroxidase, and the optical density (OD) was read at 450 nm. The binding of the residual gp120 to plate-bound D1D2 was calculated by using the following formula:

$$\text{binding (\%)} = 100 \times \frac{\text{OD}_{\text{gp120-competitor}} - \text{OD}_{\text{background}}}{\text{OD}_{\text{gp120}} - \text{OD}_{\text{background}}}$$

IC₅₀ values were obtained by nonlinear regression fitting of the binding results by Prism 4 (GraphPad Software) using the formula of one-site competition as shown below:

$$\text{binding (\%)} = \text{bottom} + \frac{\text{top} - \text{bottom}}{1 + 10^{(\log(\text{concentration}) - \log(\text{IC}_{50}))}}$$

Acknowledgment. We thank Mary Ann Gawinowicz and Yelena Milgrom for mass spectrometry and N-terminal sequencing; Joseph Sodroski and Irwin Chaiken for discussions on experimental designs; Joseph Sodroski, Shi-Hua Xiang, and Christine Winter for providing reagents; and Joel Courter for help with organizing supplemental data. This research was supported in part by National Institutes of Health Grant P01-GM 056550.

Supporting Information Available: Experimental details of the syntheses, characterization results of thiol-reacting compounds, and comprehensive information of all derivatized D1D2. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Barre-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dautet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **1983**, *220* (4599), 868–871.
- (2) Gallo, R. C.; Salahuddin, S. Z.; Popovic, M.; Shearer, G. M.; Kaplan, M.; Haynes, B. F.; Palker, T. J.; Redfield, R.; Oleske, J.; Safai, B.; White, G.; Foster, P.; Markham, P. D. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **1984**, *224* (4648), 500–503.
- (3) De Clercq, E. Emerging anti-HIV drugs. *Expert Opin. Emerging Drugs* **2005**, *10* (2), 241–273.
- (4) Klatzmann, D.; Champagne, E.; Chamaret, S.; Gruest, J.; Guetard, D.; Hecend, T.; Gluckman, J. C.; Montagnier, L. T-Lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **1984**, *312* (5996), 767–768.
- (5) Dagleish, A. G.; Beverley, P. C.; Clapham, P. R.; Crawford, D. H.; Greaves, M. F.; Weiss, R. A. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **1984**, *312* (5996), 763–767.
- (6) Trkola, A.; Dragic, T.; Arthos, J.; Binley, J. M.; Olson, W. C.; Allaway, G. P.; Cheng-Mayer, C.; Robinson, J.; Maddon, P. J.; Moore, J. P. CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. *Nature* **1996**, *384* (6605), 184–187.

- (7) Wu, L.; Gerard, N. P.; Wyatt, R.; Choe, H.; Parolin, C.; Ruffing, N.; Borsetti, A.; Cardoso, A. A.; Desjardins, E.; Newman, W.; Gerard, C.; Sodroski, J. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* **1996**, *384* (6605), 179–183.
- (8) Dimitrov, A. S.; Xiao, X.; Dimitrov, D. S.; Blumenthal, R. Early intermediates in HIV-1 envelope glycoprotein-mediated fusion triggered by CD4 and co-receptor complexes. *J. Biol. Chem.* **2001**, *276* (32), 30335–30341.
- (9) Furuta, R. A.; Wild, C. T.; Weng, Y.; Weiss, C. D. Capture of an early fusion-active conformation of HIV-1 gp41. *Nat. Struct. Biol.* **1998**, *5* (4), 276–279.
- (10) Lu, M.; Blacklow, S. C.; Kim, P. S. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat. Struct. Biol.* **1995**, *2* (12), 1075–1082.
- (11) Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **1997**, *387* (6631), 426–430.
- (12) Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **1997**, *89* (2), 263–273.
- (13) Kwong, P. D.; Wyatt, R.; Majeed, S.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates. *Structure* **2000**, *8* (12), 1329–1339.
- (14) Kwong, P. D.; Wyatt, R.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **1998**, *393* (6686), 648–659.
- (15) Huang, C. C.; Tang, M.; Zhang, M. Y.; Majeed, S.; Montabana, E.; Stanfield, R. L.; Dimitrov, D. S.; Korber, B.; Sodroski, J.; Wilson, I. A.; Wyatt, R.; Kwong, P. D. Structure of a V3-containing HIV-1 gp120 core. *Science* **2005**, *310* (5750), 1025–1028.
- (16) Wyatt, R.; Kwong, P. D.; Desjardins, E.; Sweet, R. W.; Robinson, J.; Hendrickson, W. A.; Sodroski, J. G. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **1998**, *393* (6686), 705–711.
- (17) Sattentau, Q. J.; Moore, J. P.; Vignaux, F.; Traincard, F.; Poignard, P. Conformational changes induced in the envelope glycoproteins of the human and simian immunodeficiency viruses by soluble receptor binding. *J. Virol.* **1993**, *67* (12), 7383–7393.
- (18) Thali, M.; Moore, J. P.; Furman, C.; Charles, M.; Ho, D. D.; Robinson, J.; Sodroski, J. Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120–CD4 binding. *J. Virol.* **1993**, *67* (7), 3978–3988.
- (19) Thali, M.; Furman, C.; Ho, D. D.; Robinson, J.; Tilley, S.; Pinter, A.; Sodroski, J. Discontinuous, conserved neutralization epitopes overlapping the CD4-binding region of human immunodeficiency virus type 1 gp120 envelope glycoprotein. *J. Virol.* **1992**, *66* (9), 5635–5641.
- (20) Xiang, S. H.; Kwong, P. D.; Gupta, R.; Rizzuto, C. D.; Casper, D. J.; Wyatt, R.; Wang, L.; Hendrickson, W. A.; Doyle, M. L.; Sodroski, J. Mutagenic stabilization and/or disruption of a CD4-bound state reveals distinct conformations of the human immunodeficiency virus type 1 gp120 envelope glycoprotein. *J. Virol.* **2002**, *76* (19), 9888–9899.
- (21) Chen, B.; Vogan, E. M.; Gong, H.; Skehel, J. J.; Wiley, D. C.; Harrison, S. C. Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* **2005**, *433* (7028), 834–841.
- (22) Zhou, T.; Xu, L.; Dey, B.; Hessel, A. J.; Van Ryk, D.; Xiang, S. H.; Yang, X.; Zhang, M. Y.; Zwick, M. B.; Arthos, J.; Burton, D. R.; Dimitrov, D. S.; Sodroski, J.; Wyatt, R.; Nabel, G. J.; Kwong, P. D. Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **2007**, *445* (7129), 732–737.
- (23) Myszka, D. G.; Sweet, R. W.; Hensley, P.; Brigham-Burke, M.; Kwong, P. D.; Hendrickson, W. A.; Wyatt, R.; Sodroski, J.; Doyle, M. L. Energetics of the HIV gp120–CD4 binding reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (16), 9026–9031.
- (24) Yuan, W.; Bazick, J.; Sodroski, J. Characterization of the multiple conformational states of free monomeric and trimeric human immunodeficiency virus envelope glycoproteins after fixation by cross-linker. *J. Virol.* **2006**, *80* (14), 6725–6737.
- (25) Vermeire, K.; Schols, D. Anti-HIV agents targeting the interaction of gp120 with the cellular CD4 receptor. *Expert Opin. Invest. Drugs* **2005**, *14* (10), 1199–1212.
- (26) Lin, P. F.; Blair, W.; Wang, T.; Spicer, T.; Guo, Q.; Zhou, N.; Gong, Y. F.; Wang, H. G.; Rose, R.; Yamanaka, G.; Robinson, B.; Li, C. B.; Fridell, R.; Deminie, C.; Demers, G.; Yang, Z.; Zadjura, L.; Meanwell, N.; Colonna, R. A small molecule HIV-1 inhibitor that targets the HIV-1 envelope and inhibits CD4 receptor binding. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (19), 11013–11018.
- (27) Guo, Q.; Ho, H. T.; Dicker, I.; Fan, L.; Zhou, N.; Friberg, J.; Wang, T.; McAuliffe, B. V.; Wang, H. G.; Rose, R. E.; Fang, H.; Scarnati, H. T.; Langley, D. R.; Meanwell, N. A.; Abraham, R.; Colonna, R. J.; Lin, P. F. Biochemical and genetic characterizations of a novel human immunodeficiency virus type 1 inhibitor that blocks gp120–CD4 interactions. *J. Virol.* **2003**, *77* (19), 10528–10536.
- (28) Wang, T.; Zhang, Z.; Wallace, O. B.; Deshpande, M.; Fang, H.; Yang, Z.; Zadjura, L. M.; Tweedie, D. L.; Huang, S.; Zhao, F.; Ranadive, S.; Robinson, B. S.; Gong, Y. F.; Ricarrdi, K.; Spicer, T. P.; Deminie, C.; Rose, R.; Wang, H. G.; Blair, W. S.; Shi, P. Y.; Lin, P. F.; Colonna, R. J.; Meanwell, N. A. Discovery of 4-benzoyl-1-[(4-methoxy-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)oxoacetyl]-2-(*R*)-methylpiperazine (BMS-378806): a novel HIV-1 attachment inhibitor that interferes with CD4-gp120 interactions. *J. Med. Chem.* **2003**, *46* (20), 4236–4239.
- (29) Ho, H. T.; Fan, L.; Nowicka-Sans, B.; McAuliffe, B.; Li, C. B.; Yamanaka, G.; Zhou, N.; Fang, H.; Dicker, I.; Dalterio, R.; Gong, Y. F.; Wang, T.; Yin, Z.; Ueda, Y.; Matisckella, J.; Kadow, J.; Clapham, P.; Robinson, J.; Colonna, R.; Lin, P. F. Envelope conformational changes induced by human immunodeficiency virus type 1 attachment inhibitors prevent CD4 binding and downstream entry events. *J. Virol.* **2006**, *80* (8), 4017–4025.
- (30) Zhao, Q.; Ma, L.; Jiang, S.; Lu, H.; Liu, S.; He, Y.; Strick, N.; Neamati, N.; Debnath, A. K. Identification of *N*-phenyl-*N'*-(2,2,6,6-tetramethyl-piperidin-4-yl)-oxalamides as a new class of HIV-1 entry inhibitors that prevent gp120 binding to CD4. *Virology* **2005**, *339* (2), 213–225.
- (31) Si, Z.; Madani, N.; Cox, J. M.; Chruma, J. J.; Klein, J. C.; Schon, A.; Phan, N.; Wang, L.; Biorn, A. C.; Cocklin, S.; Chaiken, I.; Freire, E.; Smith, A. B., III; Sodroski, J. G. Small-molecule inhibitors of HIV-1 entry block receptor-induced conformational changes in the viral envelope glycoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (14), 5036–5041.
- (32) Madani, N.; Perdigoto, A. L.; Srinivasan, K.; Cox, J. M.; Chruma, J. J.; LaLonde, J.; Head, M.; Smith, A. B., III; Sodroski, J. G. Localized changes in the gp120 envelope glycoprotein confer resistance to human immunodeficiency virus entry inhibitors BMS-806 and #155. *J. Virol.* **2004**, *78* (7), 3742–3752.
- (33) Schon, A.; Madani, N.; Klein, J. C.; Hubicki, A.; Ng, D.; Yang, X.; Smith, A. B., III; Sodroski, J.; Freire, E. Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry* **2006**, *45* (36), 10973–10980.
- (34) Cochran, A. G. Antagonists of protein–protein interactions. *Chem. Biol.* **2000**, *7* (4), R85–R94.
- (35) Huang, C. C.; Stricher, F.; Martin, L.; Decker, J. M.; Majeed, S.; Barthe, P.; Hendrickson, W. A.; Robinson, J.; Roumestand, C.; Sodroski, J.; Wyatt, R.; Shaw, G. M.; Vita, C.; Kwong, P. D. Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth. *Structure* **2005**, *13* (5), 755–768.
- (36) Bohacek, R. S.; McMartin, C. Multiple highly diverse structures complementary to enzyme binding sites: results of extensive application of a de novo design method incorporating combinatorial growth. *J. Am. Chem. Soc.* **1994**, *116* (13), 5560–5571.
- (37) Ripka, A. S.; Satyshur, K. A.; Bohacek, R. S.; Rich, D. H. Aspartic protease inhibitors designed from computer-generated templates bind as predicted. *Org. Lett.* **2001**, *3* (15), 2309–2312.
- (38) Arthos, J.; Deen, K. C.; Chaikin, M. A.; Fornwald, J. A.; Sathe, G.; Sattentau, Q. J.; Clapham, P. R.; Weiss, R. A.; McDougal, J. S.; Pietropaolo, C.; Axel, R.; Truneh, A.; Maddon, P. J.; Sweet, R. W. Identification of the residues in human CD4 critical for the binding of HIV. *Cell* **1989**, *57* (3), 469–481.
- (39) Smith, D. H.; Byrn, R. A.; Marsters, S. A.; Gregory, T.; Groopman, J. E.; Capon, D. J. Blocking of HIV-1 infectivity by a soluble, secreted form of the CD4 antigen. *Science* **1987**, *238* (4834), 1704–1707.
- (40) Gadek, T. R. Strategies and methods in the identification of antagonists of protein–protein interactions. *BioTechniques* **2003**, Suppl., 21–24.
- (41) Carr, R.; Jhoti, H. Structure-based screening of low-affinity compounds. *Drug Discovery Today* **2002**, *7* (9), 522–527.
- (42) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (17), 9367–9372.
- (43) Smith, A. B., III; Savinov, S. N.; Manjappara, U. V.; Chaiken, I. M. Peptide–small molecule hybrids via orthogonal deprotection–chemoselective conjugation to cysteine-anchored scaffolds. A model study. *Org. Lett.* **2002**, *4* (23), 4041–4044.
- (44) Martin, L.; Stricher, F.; Misse, D.; Sironi, F.; Pugniere, M.; Barthe, P.; Prado-Gotor, R.; Freulon, I.; Magne, X.; Roumestand, C.; Menez, A.; Lusso, P.; Veas, F.; Vita, C. Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes. *Nat. Biotechnol.* **2003**, *21* (1), 71–76.

- (45) Allaway, G. P.; Davis-Bruno, K. L.; Beaudry, G. A.; Garcia, E. B.; Wong, E. L.; Ryder, A. M.; Hasel, K. W.; Gauduin, M. C.; Koup, R. A.; McDougal, J. S.; Maddon, P. J. Expression and characterization of CD4-IgG2, a novel heterotetramer that neutralizes primary HIV type 1 isolates. *AIDS Res. Hum. Retroviruses* **1995**, *11* (5), 533–539.
- (46) Gauduin, M. C.; Allaway, G. P.; Maddon, P. J.; Barbas, C. F., 3rd; Burton, D. R.; Koup, R. A. Effective ex vivo neutralization of human immunodeficiency virus type 1 in plasma by recombinant immunoglobulin molecules. *J. Virol.* **1996**, *70* (4), 2586–2592.
- (47) Trkola, A.; Pomales, A. B.; Yuan, H.; Korber, B.; Maddon, P. J.; Allaway, G. P.; Katinger, H.; Barbas, C. F., 3rd; Burton, D. R.; Ho, D. D.; et al. Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG. *J. Virol.* **1995**, *69* (11), 6609–6617.
- (48) Arthos, J.; Cicala, C.; Steenbeke, T. D.; Chun, T. W.; Dela Cruz, C.; Hanback, D. B.; Khazanie, P.; Nam, D.; Schuck, P.; Selig, S. M.; Van Ryk, D.; Chaikin, M. A.; Fauci, A. S. Biochemical and biological characterization of a dodecameric CD4–Ig fusion protein: implications for therapeutic and vaccine strategies. *J. Biol. Chem.* **2002**, *277* (13), 11456–11464.
- (49) Culp, J. S.; Johansen, H.; Hellmig, B.; Beck, J.; Matthews, T. J.; Delers, A.; Rosenberg, M. Regulated expression allows high level production and secretion of HIV-1 gp120 envelope glycoprotein in *Drosophila Schneider* cells. *Biotechnology (N. Y.)* **1991**, *9* (2), 173–177.
- (50) Dey, B.; Pancera, M.; Svehla, K.; Shu, Y.; Xiang, S. H.; Vainshtein, J.; Li, Y.; Sodroski, J.; Kwong, P. D.; Mascola, J. R.; Wyatt, R. Characterization of human immunodeficiency virus type 1 monomeric and trimeric gp120 glycoproteins stabilized in the CD4-bound state: antigenicity, biophysics, and immunogenicity. *J. Virol.* **2007**, *81* (11), 5579–5593.
- (51) Korber, B.; Foley, F.; Kuiken, C.; Pillai, S.; Sodroski, J. Numbering Positions in HIV Relative to HXBc2. In *Human Retroviruses and AIDS*; Korber, B., Kuiken, C. L.; Foley, B., Hahn, B., McCutchan, F., Mellors, J. W., Sodroski, J., Eds.; Los Alamos National Laboratory: Los Alamos, NM, 1998; pp III-102 through IV-103.
- (52) Osburne, M. S.; Neidhardt, E. A.; Godoy, J. E.; van Schravendijk, M. R.; Grossman, T. H. Production of secreted, soluble human two-domain CD4 protein in *Escherichia coli*. *J. Immunol. Methods* **1999**, *224* (1–2), 19–24.
- (53) Coligan, J. E.; Dunn, B. M.; Ploegh, H. L.; Speicher, D. W.; Wingfield, P. T. *Current Protocols in Protein Science*; John Wiley & Sons: 2001.

JM070564E