Protein Minimization of the gp120 Binding Region of Human CD4[†]

Deepak Sharma,[‡] M. M. Balamurali,[‡] Kausik Chakraborty,[‡] Sowmini Kumaran,[‡] Sadasivam Jeganathan,[‡] Umar Rashid,[‡] Paolo Ingallinella,[§] and Raghavan Varadarajan*,[‡]

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, and Istituto di Ricerche di Biologia Molecolare P. Angeletti, Via Pontina Km 30.600, 00040 Pomezia, Italy

Received June 11, 2005; Revised Manuscript Received September 14, 2005

ABSTRACT: CD4 is an important component of the immune system and is also the cellular receptor for HIV-1. CD4 consists of a cytoplasmic tail, one transmembrane region, and four extracellular domains, D1−D4. Constructs consisting of all four extracellular domains of human CD4 as well as the first two domains (CD4D12) have previously been expressed and characterized. All of the gp120-binding residues are located within the first N-terminal domain (D1) of CD4. To date, it has not been possible to obtain domain D1 alone in a soluble and active form. Most residues in CD4 that interact with gp120 lie within the region 21−64 of domain D1 of CD4. On the basis of these observations and analysis of the crystal structure of CD4D12, a mutational strategy was designed to express CD4D1 and region 21−64 of CD4 (CD4PEP1) in *Escherichia coli*. K_D values for the binding of CD4 analogues described above to gp120 were measured using a Biacore-based solution-phase competition binding assay. Measured K_D values were 15 nM, 40 nM, and 26 μM for CD4D12, CD4D1, and CD4PEP1, respectively. All of the proteins interact with gp120 and are able to expose the 17b-binding epitope of gp120. Structural content was determined using CD and proteolysis. Both CD4D1 and CD4PEP1 were partially structured and showed an enhanced structure in the presence of the osmolyte sarcosine. The aggregation behavior of all of the proteins was characterized. While CD4D1 and CD4PEP1 did not aggregate, CD4D12 formed amyloid fibrils at neutral pH within a week at 278 K. These CD4 derivatives should be useful tools in HIV vaccine design and entry inhibition studies.

CD4 is a 55-kDa glycoprotein expressed on approximately 60% of the peripheral blood T lymphocytes (*1*). CD4 is also expressed at lower levels on macrophages, dendritic cells, Langerhans cells, hematopoietic stem cells, and microglial cells. CD4 consists of four extracellular domains (residues 1−371), a transmembrane segment (residues 372−395), and a short cytoplasmic tail (residues 396−433). It activates T lymphocytes by binding to the nonpolymorphic region of the major histocompatibility complex (MHC) class-II antigens expressed on the surface of antigen-presenting cells (*2*). CD4 is a very important component of the immune system, and new functional roles of CD4 are still being discovered (*3*, *4*). In humans, CD4 is the primary receptor of HIV (human immunodeficiency virus) (*5*), binding with high affinity to the viral envelope glycoprotein gp120 (*6*). It is required for viral attachment and subsequent entry into host cells. The crystal structures of the entire extracellular four domains of human CD4 (*7*), the two N-terminal domains of human CD4 (hereafter referred to as CD4D12)[1] (*8*−*10*), and complexes of CD4D12 with class-II MHC (*11*) as well as with the HIV envelope protein gp120 (*12*) are known.

HIV entry into T cells is mediated by the binding of the gp120 subunit of Env with CD4. The interaction of gp120 with CD4 leads to conformational changes in the Env protein and its subsequent binding with coreceptor CCR5 and/or CXCR4. Although viral isolates that do not require CD4 for entry into host cells have been identified (*13*), their clinical significance is uncertain. There have been several attempts to use rsCD4 as an inhibitor of HIV binding to gp120 (*14*, *15*). The CD4-binding epitope of gp120 is highly conserved among different viral isolates and therefore is an important target for both therapeutic strategies and vaccination. It has been shown that, *in vitro*, the rsCD4 protein efficiently inhibits virus infection and virus-mediated cell fusion (*16*). Unfortunately, administration of rsCD4 *in vivo* does not block infection of primary HIV-1 isolates. This is in part because of the short serum half-life of rsCD4 (*17*) (approximately 30 min). It was also found that a 1000-fold higher concentration of recombinant soluble four-domain CD4 (rsCD4) was required to neutralize primary isolates

[†] This work was supported by grants from the NIH AIDS Innovation grant program (NIH-AI-46984) and Department of Science and Technology, Government of India, to R.V.. R.V. is a recipient of a Senior Research Fellowship from the Wellcome Trust.

* To whom correspondence should be addressed: Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India. E-mail: varadar@mbu.iisc.ernet.in. Telephone: 91-80-22932612. Fax: 91-80-23600535.

[‡] Indian Institute of Science.

[§] Istituto di Ricerche di Biologia Molecolare P. Angeletti.

[1] Abbreviations: CD, circular dichroism; ELISA, enzyme-linked immunosorbent assay; HPLC, high-pressure liquid chromatography; GdnCl, guanidinium chloride; PBS, phosphate-buffered saline; CD4D12, two N-terminal domains of human CD4 (amino acids 1−183); CD4D1, first N-terminal domain of human CD4 (amino acids 1−99); rsCD4 recombinant soluble four domain CD4; hCD4, human CD4; N-HisTrx, N-terminal histidine-tagged *E. coli* thioredoxin.

10.1021/bi051120s CCC: $30.25
Published on Web 11/12/2005

relative to tissue-culture-adapted strains (*18*) and that lower concentrations of rsCD4 enhance the infectivity of primary isolates (*19*). This is of primary concern because of the short serum half-life of rsCD4. Recently, cross-linked complexes of rsCD4 with gp120 were used as immunogens in rhesus macaques (*20*). The resulting sera were able to neutralize diverse strains of HIV-1. In a subsequent study, single-chain derivatives of gp120 linked to the first two domains of CD4 were used as immunogens in guinea pigs (*21*). The resulting sera showed broad neutralization. However, it was shown that this was primarily due to the anti-CD4 antibodies. Because only residues in a restricted region of CD4D1 are involved in contacting gp120, it was of interest to design a smaller version of CD4 with similar gp120-binding affinity to CD4D12, to minimize the immune response against CD4 in the corresponding single-chain analogues. Other potential applications of such constructs would be in the inhibition of viral entry and also as reagents to help elucidate the functions of the various domains of CD4. Recently, lactobacilli expressing soluble CD4 have been shown to inhibit HIV entry and are being considered for microbicidal use (*22*). We have therefore attempted to produce and characterize various soluble derivatives of the gp120-binding region of CD4.

Structural as well as sequence alignments show that CD4 belongs to the immunoglobulin superfamily (IgSF). Only domain 1 of CD4 (CD4D1) makes direct contact with the class-II MHC molecule or with gp120. Most of the residues important for binding to gp120 lie between residues 25−85 of domain 1. Isolated V domains of antibodies have been expressed and well-characterized in terms of their stability and folding behavior (*23*). However, thus far, it has not been possible to express and isolate CD4D1 to homogeneity. An earlier approach to produce secreted CD4−immunoglubulin chimeras failed with D1 or D3 fused to C_κ (immunoglobulin constant region of mouse), whereas D1.D2 or D1.D2.D3.D4 fused to C_κ was well-secreted, suggesting that the isolated D1 domain is unstable in the absence of D2 (*24*, *25*). This is probably because domains D1 and D2 pack closely against one another with a large hydrophobic interface (700−800 Å^2). The last strand of D1 continues into the first strand of D2 with residue 98 forming a hydrogen bond within a sheet of D1 and residue 99 forming a hydrogen bond in a sheet of D2. This tight connectivity between domains 1 and 2 (*26*) presumably makes it difficult to express well-folded CD4D1 alone. Domains 1 and 2 are thought to be involved in the interaction with MHC II molecules (*27*), although the crystal structure suggests that only D1 is involved in the physical interaction with MHC II. D3 and D4 domains possibly play a role in the interaction of CD4 with the T-cell receptor (*28*). The cytoplasmic domain of CD4 is noncovalently associated with the protein tyrosine kinase p56^{lck} (*29*).

In this study, we report the rational design and characterization of the following derivatives of soluble CD4: (1) CD4D12 consisting of amino acids 1−183 of human CD4 (hCD4), (2) CD4D1 (amino acids 1−99 of hCD4), (3) CD4PEP1 (amino acids 21−64 of hCD4), and (4) Trx-CD4PEP1 having CD4PEP1 inserted in one of the surface loops of *Escherichia coli*, thioredoxin. All constructs were expressed in *E. coli* and purified to homogeneity. The aggregation behavior of all of the constructs was characterized. CD4D12 alone was found to form amyloid-like fibrils at neutral pH within 5−6 days at room temperature. The binding affinity of each CD4 derivative for gp120 was measured using surface plasmon resonance. In all cases, binding was shown to expose the epitope on gp120 for binding the monoclonal antibody 17b.

MATERIALS AND METHODS

DNA Sequencing and Mass Spectrometry. The identities of all of the mutants and constructs described below were confirmed by DNA sequencing. In all cases, the entire coding region was sequenced. ESI−MS was carried out using an esquire$_{3000}^{plus}$ instrument from Bruker-Daltonics.

Construct Descriptions. CD4D12 refers to the first two N-terminal domains (amino acids 1−183) of hCD4. D1 refers to domain 1 (amino acids 1−98) and D2 refers to domain 2 (amino acids 98−183) of hCD4. CD4D1 (amino acids 1−99) is an engineered version of D1 with additional mutations and an additional residue at the C terminus. CD4PEP1 encodes residues from 21 to 64 of hCD4 containing several mutations. TrxCD4PEP1 contains CD4PEP1 inserted in one of the exposed loops of *E. coli* thioredoxin.

Construction and Expression of CD4D12 and CD4D1. CD4D12 was expressed and purified as described previously (*21*). After purification, the protein was dialyzed against 5% acetic acid and then against a buffer containing 10 mM each of citrate, glycine, and HEPES (CGH10) at pH 4.0 and stored at −70 °C. CD4D1 was obtained from CD4D12 as follows. The following mutations were introduced: V3T, L5A, I76T, L96A, and F98A. The rationale for these choices is described in a later section. Mutagenesis was done by overlap PCR. Wild-type D1 without additional mutations was also cloned into the pET9a vector. However, the protein expressed from this construct was highly aggregation-prone and could not be refolded. CD4D1 was expressed and purified using a procedure similar to that described above for CD4D12. The approximate purified yields of CD4D1 was 7 mg/L.

Construction and Expression of CD4PEP1 and Trx-CD4PEP1. CD4PEP1 was constructed and cloned using a procedure similar to that used for CD4D1. The following mutations were introduced: I24T, F26A, W28A, and L51A. CD4PEP1 was initially cloned into pET9a and subsequently into pMMHa (*30*) to generate pET9a-CD4PEP1 and pMM-CD4PEP1, respectively. CD4PEP1 was also inserted between codons 74 and 75 of His-tagged *E. coli* thioredoxin using overlap PCR in the construct pBAD-NhisTrxCD4PEP1.

TrxCD4PEP1 was expressed from the soluble fraction using the pBAD-NhisTrxCD4PEP1 vector in the *E. coli* strain DH5α. In this construct, amino acids 21−64 of CD4 have been inserted between amino acids 74 and 75 of *E. coli* thioredoxin. The construct also contains an N-terminal histidine tag to facilitate purification. After Ni-NTA affinity purification, the sample was extensively dialyzed against phosphate-buffered saline (PBS) to remove imidazole. The protein was found to be 80% pure. Further purification (>95%) was achieved using a Resource Q anion-exchange column. From the insoluble fraction, TrxCD4PEP1 fusion protein was purified using essentially the same procedure as described for CD4D12. The total yield of the purified protein was 3 mg/L from the soluble fraction and 3−4 mg/L from the insoluble fraction.

CD4PEP1 from plasmid pET9a-CD4PEP1 was purified from the soluble fraction using a procedure similar to that

described above for TrxCD4PEP1. Resource Q purification was not carried out as the peptide purified from Ni-NTA, and subsequent dialysis was approximately 90% pure. The yield of the purified peptide from this system was approximately 0.2−0.4 mg/L.

The His-tagged TrpE-CD4PEP1 fusion protein expressed from plasmid pMMCD4PEP1 was purified under denaturing conditions using Ni-NTA chromatography. The eluted sample was extensively dialyzed against 5% acetic acid. The dialyzed sample was lyophilized. CNBr cleavage [1:2 (w/w) ratio of sample/CNBr] was performed in 70% formic acid in the dark at room temperature for 12 h. The CNBr-cleaved sample was again lyophilized and then dissolved in 5% acetic acid to remove denaturant and imidazole. The sample was then extensively dialyzed against double-distilled water. The dialyzed sample was then collected, and pH was increased to 7.4 using 10× PBS to precipitate the cleaved leader sequence. The solution was spun at 9000*g* at 4 °C for 30 min. The supernatant was collected, and the purity of the peptide was characterized using SDS−PAGE and mass spectrometry. The concentration of peptide was estimated by absorbance at 280 nm using an extinction coefficient of 5500 M^{-1} cm^{-1} calculated from the amino acid sequence (*31*). The peptide was found to be >99% pure. The total yield of the purified peptide from this system was around 100−120 mg/L.

Synthesis and Purification of M33 and M9. CD4M9 (M9) is derived from a short scorpion toxin, scyllatoxin, and reproduces the gp120-binding region of CD4 (*32*). CD4M33 (M33) is derived from CD4M9 and is known to bind to gp120 with a higher affinity than M9 (*33*). M33 and M9 were synthesized as described previously (*32*, *33*). Peptides were dissolved in aqueous acetonitrile and diluted to 0.1 mg/mL in Tris buffer at pH 8.0. Peptides were oxidized in the presence of GSH/GSSG and purified using reverse-phase high-pressure liquid chromatography (HPLC) on a C18 reverse-phase column. The final purified peptides were >95% pure. Mass spectrometry confirmed that all Cys residues were oxidized to disulfides in the final purified product.

Gel-Filtration Analysis of CD4D12, CD4D1, TrxCD4PEP1, and CD4PEP1. Approximately 50−150 μg of protein was analyzed under nondenaturing conditions by gel-filtration chromatography in PBS buffer using a Superdex-75 analytical gel-filtration column. A standard curve of the elution volume versus the log of the molecular weight was generated using a set of protein standards. This curve was then used to calculate the apparent molecular weight of the above-listed proteins.

Biacore Experiments. Biacore experiments were performed with a Biacore2000 (Biacore, Uppsala, Sweden) optical biosensor at 25 °C and pH 7.4. The 17b-binding assay for gp120 complexed to CD4 analogues was performed as reported previously (*21*). The affinity of the various analogues for gp120 was quantitated using a competition binding assay described below.

The binding phase of a sensorgram can be described by the equation (BIAtechnology handbook, edition June 1994)

$$dR/dt = k_a[A]R_{max} - (k_a[A] + k_d)R \quad (1)$$

where [A] is the concentration of the analyte, R_{max} is the maximum RU of A that the surface-immobilized ligand (B) can bind, R is the response in RU that is proportional to the concentration of AB, k_a is the association rate constant, and k_d is the dissociation rate constant. R_{max}, the maximum value of R at saturation, is a fixed parameter that depends upon the amount of B immobilized and molecular weight of A. If the same chip surface with a fixed amount of B is used to examine the binding of different A values that have similar molecular weights, then R_{max} is essentially a constant.

In the initial phase of the binding curve, R is negligible compared to R_{max} (typically 2−3% of R_{max} in experiments reported here), and hence, the above equation can be approximated to $dR/dt = r_{init} \approx k[A]$, where k is the product of k_a and R_{max} and is a constant when various different A values have the same value of k_a and similar molecular weights and the same chip surface is used.

The initial 30 s of the binding phase of a sensorgram can be fit to a straight line to obtain the rate of dR/dt in the initial phase, which is henceforth referred to as r_{init}. Because $r_{init} = k[A]$, k can be obtained by a linear fit of r_{init} as a function of [A]. k obtained in this way can then be used to calculate [A] in a sample of unknown concentration from the measured value of r_{init} ([A] = r_{init}/k).

Principle and Methods of the F105 Competition Binding Assay. In the present study, the CD4-binding-site antibody F105 (*34*) was used to measure the K_D of the binding of CD4D12 and other CD4 analogues to gp120. F105 binds to the CD4-binding site and hence out competes CD4 binding to gp120. In this assay, the concentration of free gp120 in a solution containing a mixture of gp120 and CD4D12/CD4 analogues was determined by injecting it over a F105 surface. In this assay, 1000 RU of F105 antibody was immobilized on a CM5 research-grade chip with amine-coupling chemistry. Different concentrations of gp120 (50 nM, 25 nM, 12.5 nM, and 6.25 nM) were passed over the F105 surface to obtain binding sensorgrams. The first 30 s of the binding phase was fitted to a straight line to obtain the r_{init} values as a function of the gp120 concentration. These r_{init} values were plotted against the gp120 concentration and fit to a straight line to obtain k. Next, 50 nM gp120 was incubated with different concentrations of CD4 analogues and injected over the same F105 surface. Because, both F105 and CD4D12 bind at the same site on gp120, only free gp120 will bind to the F105 surface. The dissociation rate constant of CD4D12 bound to gp120 is approximately 10^{-4} s^{-1}, and the residence time of the solution over the chip is 0.1 s; therefore, no dissociation of the complexed gp120 occurs in the time scale of the experiments. Hence, the amount of free gp120 present in a mixture of free and complexed gp120 could be measured from the values of r_{init} obtained as function of the added CD4 analogue and the known value of k. The concentration of free gp120 was then used to calculate the K_D for the binding to gp120 of the CD4 analogue using the equation

$$K_D = ([gp120_{free}])([L]_T - ([gp120]_T - [gp120_{free}]))/([gp120]_T - [gp120_{free}]) \quad (2)$$

where $[gp120_{free}]$ is the concentration of free gp120 obtained from the measured value of r_{init} and the k value. $[gp120]_T$ is the total concentration of gp120 in the solution (50 nM), and $[L]_T$ is the total concentration of the CD4 analogue in the solution. It was found that the K_D values for binding were

quite reproducible at concentrations of the CD4 analogue where the inhibition was 40−60%. The K_D at very low and very high concentrations of the CD4 analogues were not reliable and hence were not used for calculating the average K_D.

Congo Red and Thioflavin T Binding Assays. The binding of congo red to proteins was monitored using absorption spectroscopy. A 5.0 mM stock of congo red was prepared in PBS. Congo red (5 μM) spectra were obtained in the presence or absence of CD4D12 suspension (10 μM). Samples were incubated for 30 min before the absorption measurement.

A total of 25 μM thioflavin T was incubated in the presence and absence of 10 μM protein solution in PBS. Thioflavin T fluorescence was recorded from 460 to 500 nm with an excitation at 450 nm and slits of 3.5 nm for both excitation and emission.

Fluorescence and Circular Dichroism (CD) Measurements. Fluorescence spectra were recorded at 25 °C on a SPEX Fluoromax3 spectrofluorimeter. For intrinsic fluorescence measurements, the excitation was at 280 nm and emission was recorded from 300 to 400 nm. The excitation and emission slit widths were 3 and 5 nm, respectively. ANS fluorescence was measured with excitation at 365 nm and emission from 400 to 600 nm. All experiments were carried out in PBS at pH 7.4. Far-UV CD measurements were performed on homogeneous protein samples at a concentration of 15 μM in 0.1 cm path-length cuvettes as described previously (*21*). Photomultiplier voltage did not exceed 800 V in the spectral region analyzed. Data were corrected for buffer contributions. The results are expressed as mean residue ellipticity (MRE) calculated according to the procedure described earlier (*21*). The population of folded CD4D1 molecules in aqueous solution (fN) was estimated as described previously (*35*) by assuming that the MRE corresponding to 100% folded CD4D1 was identical to that of CD4D12 in aqueous solution. Therefore

$$\text{fN}\ (\%) = \frac{[\theta]_{\text{CD4D1,water}} - [\theta]_{\text{CD4D1,GdnCl}}}{[\theta]_{\text{CD4D12,water}} - [\theta]_{\text{CD4D12,GdnCl}}} \times 100 \quad (3)$$

where $[\theta]_{\text{CD4D1,water}}$ is the MRE of CD4D1 at a given wavelength in aqueous solution, $[\theta]_{\text{CD4D1,GdnCl}}$ is the corresponding MRE in 6 M guanidinium chloride (GdnCl), and $[\theta]_{\text{CD4D1,water}}$ is the corresponding MRE of CD4D12 in aqueous solution.

For denaturation studies of CD4D12 and CD4D1, 8 μM protein was incubated in increasing concentrations of GdnCl in PBS overnight at 25 °C. The CD signal was measured at 215 nm in a 0.1 cm path-length cuvette to monitor the extent of denaturation.

Proteolysis. Proteolytic digestion of CD4 analogues in PBS were carried out using trypsin at a protease/substrate ratio of 1:50 (w/w). A total of 15−25 μg of protein was digested in 200 μL of buffer. Digestion was carried out both in the presence and absence of sarcosine. We have previously shown that the specific activity of trypsin is lowered by a factor of 2 in the presence of 6 M sarcosine (*36*), and hence, a 2-fold higher concentration of trypsin was used in the presence of 6 M sarcosine. At various times, 50 μL of sample was removed and was immediately boiled in SDS−PAGE gel-loading buffer (50 mM Tris-HCl at pH 6.8 containing 2.0% SDS, 0.1% bromophenol blue, and 5% β-mercaptoethanol) for 10 min and stored at −20 °C. Samples collected at different time points were subjected together to SDS−PAGE electrophoresis followed by staining with Coomassie Brilliant Blue R250.

Electron Microscopy. CD4D12 was stained on carbon-coated copper grids. Samples were adsorbed for up to 5 min, briefly washed, and stained with a freshly filtered 2% uranyl acetate solution. After drying, samples were viewed in an electron microscope (Geol 100CXII; voltage, 80 000 V; magnification, 40 000).

RESULTS

Design of CD4D1. Domains D1 and D2 pack against one another with a large hydrophobic interface (700−800 Å^2). The last strand of D1 continues straight into D2 with residue 98 forming a hydrogen bond with a strand of D1 and residue 99 forming a hydrogen bond with a sheet of D2. Several mutations in the D2 domain of CD4 have been shown to interfere with CD4 binding to gp120 (*37*). However, all of the residues directly interacting with gp120 appear to lie in domain D1 of CD4 (*38*). These data suggest that there is a strong interdomain interaction between the D1 and D2 domains.

Using accessible surface area calculations of residues in CD4D12 described below, we identified all of the important residues in the D1 domain, which interact significantly with residues in the D2 domain of CD4 (Figure 1) as well as residues that interact with gp120. Interacting residues of CD4D12 with gp120 were identified on the basis of the change in the accessible surface area of the residues of CD4D12 in the presence and absence of gp120. Residues in CD4D12 having an accessible surface area greater by 5 Å^2 in the absence of gp120 than in the presence of gp120 were considered to be interacting residues with gp120. Interacting residues with gp120 lie primarily in the region 25−64. However, many of the residues in the 25−64 region were found to interact with the residues lying outside this region within domain 1; therefore, expression of just this region is not feasible. According to the SCOP definition (*39*), domain D1 of CD4 comprises residues 1−97. However, we have also included residues 98 and 99 because those were found to be interacting residues with several other residues of CD4D1. Accessible surface areas (ASA) of all of the residues in D1 were calculated in the presence and absence of D2. All of the residues with a change in ASA (ΔASA = ASA in the absence of domain D2 − ASA in the presence of domain D2) of greater than 10 Å^2 were classified as the residues involved in the interaction with D2. None of the residues interacting with the D2 domain of CD4D12 also interacted with gp120. Those hydrophobic residues in D1 that interact with D2 were mutated to either Thr or Ala. On the basis of these observations and analysis, region 1−99 with the above-designed mutations was chosen as the minimum region of CD4 that might be well-folded and interact with gp120 with an affinity similar to CD4D12. CD4D1 could be expressed and refolded in high yield. An earlier study attempted to express the first 113 amino acids of hCD4. However, the yield of purified refolded protein was approximately 50-fold lower than in the present study, and no biophysical characterization was reported (*40*).

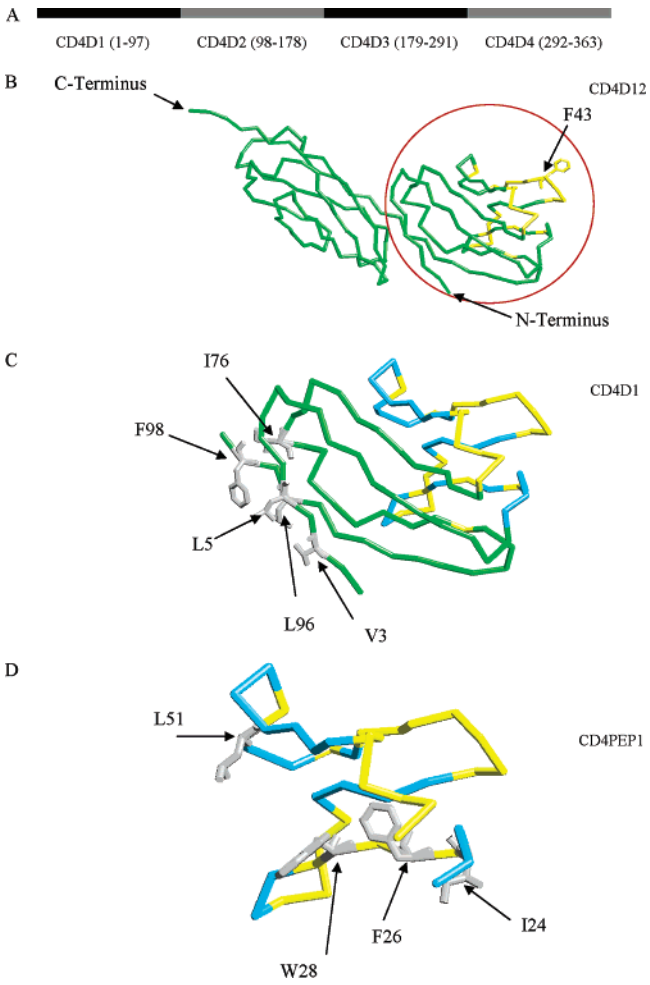

FIGURE 1: Locations of CD4D1 and CD4PEP1 in the structure of CD4D12. (A) Linear schematic shows the arrangement of four extracellular domains of hCD4. (B) Two N-terminal domains of hCD4. The region within the red circle in CD4D12 is the 99-residue N-terminal domain (CD4D1) of CD4. Yellow regions correspond to the gp120-interacting residues of CD4. The side chain of the critical gp120-binding residue F43 is also shown. (C) CD4D1 domain. The CD4PEP1 region is shown in cyan, except for gp120-interacting residues, which are in yellow. (D) CD4PEP1 fragment. Exposed hydrophobic residues mutated in either CD4D1 or CD4PEP1 are shown in gray.

Design of CD4PEP1. The crystal structure of core gp120 in a ternary complex with CD4 and 17bF_{ab} as well as mutagenesis work reported previously (*12*, *41*) show that most of the residues important for the interaction with gp120 lie in the region from 25 to 64 (25, 27, 29, 33−35, 40−46, 48, 52, and 59−64) of CD4D12. Of a total of 970 Å^2 buried upon the interaction with gp120, the 25−64 region accounts for approximately 965 Å^2. To locate residues in 25−64 that interact with residues lying outside the region 25−64, accessible surface areas of all residues 25−64 were calculated in the presence and absence of CD4D12. A cutoff of >10 Å^2 for ΔASA [ΔASA = ASA in the absence of region outside residue (25−64) − ASA in the presence of all of the residues of domain D1] was used. Residues 21, 23, and 24 were found to interact with residues 62 and 63 in the region 25−64 of CD4D12 (ΔASA ∼ 30 Å^2 for each residue). The C_α−C_α distance between residues 21 and 64 is 4.5 Å. On the basis of these observations, region 21−64 was chosen as the minimum region that could form a well-defined structure (Figure 1). All of the hydrophobic residues within region 21−64 with ΔASA > 10 Å^2 and which do not interact with gp120 (listed in the Materials and Methods) were mutated to either Thr or Ala. However, I36 and L61 were not mutated because these are adjacent to K35 and W62, respectively, which have been shown to play an important role in the binding of CD4D12 to gp120 (*41*). Designed CD4PEP1 was initially cloned in the expression vector pET9a and purified using Ni-NTA affinity purification as described in the Materials and Methods. However, the yield of the purified CD4PEP1 was too low for any further characterization. CD4PEP1 was then inserted into one of the loops of *E. coli* thioredoxin between residues 74−75. We have recently shown (Chakraborty and Varadarajan, unpublished results) that this site tolerates insertion better than the active-site region used for peptide insertion in earlier studies (*42*, *43*). Fusion to thioredoxin has been shown to produce many mammalian proteins in high yield in *E. coli*. The yield of purified fusion protein in this case was 3 mg/L of culture. We chose to insert the peptide into a loop region rather than at the C terminus of thioredoxin to bring the N and C termini of the peptide in close proximity. This should decrease the magnitude of the unfavorable conformational entropy change associated with folding. To further increase the yield, CD4PEP1 was subcloned as a fusion with the Trp leader sequence in pMMHa (*30*) and was purified as described in the Materials and Methods.

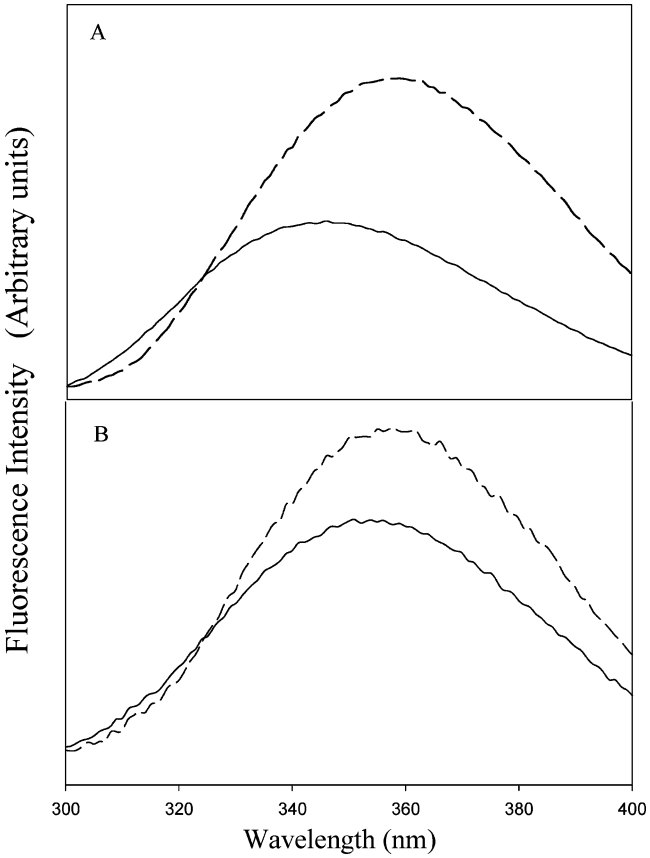

FIGURE 2: Fluorescence emission spectra for CD4D12 and CD4D1. Protein samples were incubated at a final concentration of 2 μM either in PBS (—) or in 6 M GdnCl (- - -) at 25 °C and pH 7.4 for 2 h. Spectra were obtained with excitation at 280 nm and emission from 300 to 400 nm. (A) CD4D12. (B) CD4D1.

Spectral Characterization of CD4D12 and CD4D1. Figure 2 shows the intrinsic fluorescence emission spectra of native CD4D12 in buffer and in the presence of 6 M GdnCl. Native

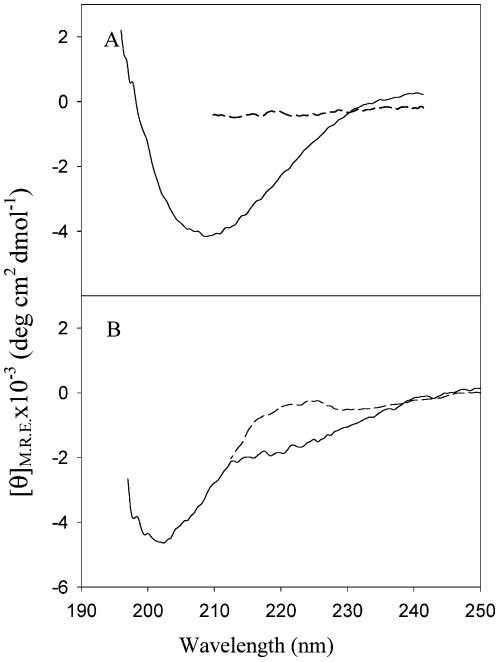

FIGURE 3: Far-UV CD spectra of CD4D12 and CD4D1. Spectra were obtained with 10 μM CD4D12 and 20 μM CD4D1 in PBS without (—) or with 6 M GdnCl at pH 7.4 (- - -) with a 0.1 cm path-length cuvette. (A) CD4D12. (B) CD4D1.

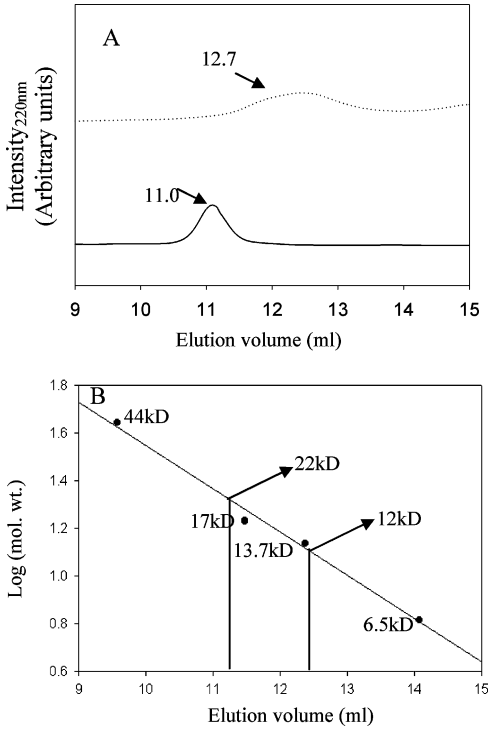

FIGURE 4: Gel-filtration analysis of CD4D12 and CD4D1 on a Superdex 75 gel-filtration column in PBS at room temperature. (A) From bottom to top: CD4D12 (15 μM) (—) and CD4D1 (50 μM) (- - -). (B) Calibration curve.

CD4D12 exhibited an emission maximum at 340 nm. The presence of 6 M GdnCl leads to a significant red shift to 360 nm and an increase in intensity of fluorescence. The shift to a longer wavelength indicates an increased exposure of tryptophan residues of CD4D12 to the aqueous medium upon denaturation. The increase in intensity upon denaturation suggests that one or more of the tryptophan residues in native CD4D12 are quenched. In the native state, Trp 28 is packed against a disulfide present in the first domain and Trp 62 is located near the guanidinium group of Arg 59. Both of these are potential quenchers (*44*, *45*). For native CD4D1, maximum emission intensity was observed at 350 nm and the presence of GdnCl leads to a red shift to 360 nm. Because the emission peak for CD4D1 is red-shifted with respect to CD4D12, this shows that either or both of the tryptophan residues at position 28 and 62 in some fraction of CD4D1 molecules are partly exposed. There was a significant increase in the fluorescence intensity and a red shift of 10 nm upon unfolding for CD4D1, indicating that a reasonable fraction of CD4D1 molecules is structured in aqueous solution.

The secondary structures of both proteins were analyzed by far-UV CD (Figure 3). CD spectra show that, under native conditions, CD4D12 has a well-defined secondary structure; however, CD4D1 shows a spectrum characteristic of proteins with short irregular β strands (*46*). The spectrum of CD4D1 differs from that of natively unfolded proteins (*47*, *48*), which have substantially more negative ellipticities ($< -15\,000°$ cm^2 dmol^{-1}) at 200 nm than CD4D1. There was a significant further decrease in the magnitude of MRE for both CD4D12 and CD4D1 when incubated in 6 M GdnCl. This indicates that a fraction of CD4D1 molecules are structured in native buffer (PBS). The fraction of folded CD4D1 was estimated to be approximately 50% using eq 3 in the Materials and Methods.

ANS-binding studies were carried out to compare the structural compactness of these CD4 derivatives. ANS is a polarity sensitive fluorescence probe and binds to hydrophobic surfaces in partially folded intermediates with much higher affinity than to the native protein, resulting in a marked increase in fluorescence emission spectra compared with the emission of free ANS in aqueous solution. For CD4D12, ANS did not bind to the native protein; however, for CD4D1, appreciable binding of ANS was observed (see curve 1 in Figure 1 of the Supporting Information).

Gel-Filtration Analysis and Denaturation Studies of CD4D12 and CD4D1. To examine the aggregation state of all of the proteins described above, size-exclusion chromatography was performed. Both molecules eluted as monomers in native conditions (PBS). Soluble CD4D12 eluted at 11.0 mL, corresponding to a molecular mass of 21 000 Da as compared with globular protein standards (Figure 4). CD4D1 eluted at 12.7 mL, corresponding to an apparent molecular mass of 11 000 Da. However, the CD4D1 peak was broader than the CD4D12 peak, suggesting that it may be composed of a mixture of folded and unfolded species in rapid equilibrium. To confirm that an appreciable fraction of CD4D1 molecules are folded at room temperature, denaturation of CD4D12 and CD4D1 was monitored by measuring the CD signal at 215 nm as a function of [GdnCl]. CD4D12 shows a transition with an apparent midpoint of 2 M. For CD4D1, no native baseline is present but the protein appears to be completely unfolded by 2 M GdnCl (Figure 5).

Amyloid Formation for CD4D12. Congo red is a diazo dye that is widely used to identify amyloid deposits because of its ability to bind preferentially to aggregated amyloid peptides/proteins but not to native unassembled forms (*49*). CD4D12 starts precipitating when stored at 4 °C at pH 7.4 after 5−6 days. When the protein is thermally denatured at

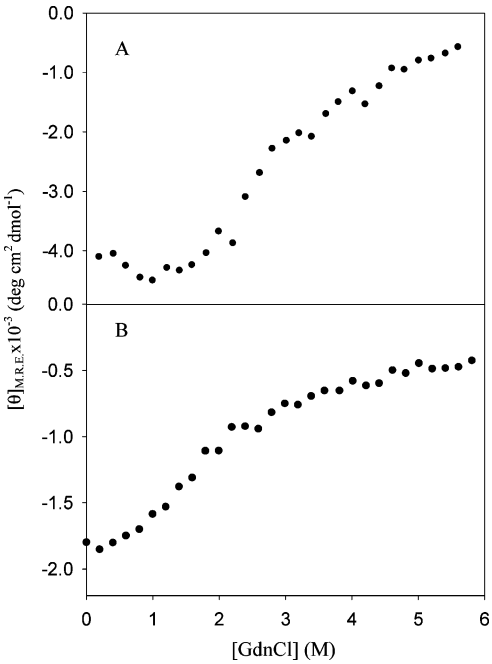

FIGURE 5: Unfolding of CD4D12 (A) and CD4D1 (B) as monitored by CD. The protein (8 μM) was incubated overnight in an increasing concentration of GdnCl at 25 °C and pH 7.4. The CD signal was measured at 215 nm in 0.1 cm path-length cuvette. Mean residue ellipticity was plotted as a function of [GdnCl].

85 °C, visible aggregates are rapidly formed. To determine whether these aggregates are amyloid in nature, a congo red binding assay was performed. Congo red binding to amyloid is detected as a red shift in its absorbance spectrum. Figure 2A in the Supporting Information shows the absorption spectrum of congo red in the presence and absence of CD4D12 aggregates along with the corresponding difference spectrum. A significant red shift from 494 to 550 nm occurred upon binding of the dye to CD4D12 for both the room temperature and thermally denatured protein.

CD4D1 did not form any visible aggregates upon storage at 4 °C or when thermally denatured. CD4D1 also did not exhibit congo red binding. Thioflavin T is a fluorescent dye that detects amyloid fibril formation (*50*). The fluorescence emission intensity of the dye increases significantly upon binding to the linear array of the β strands in amyloid fibrils. Changes in the emission intensity of thioflavin T in the presence of CD4D12 at pH 7.4 and at 25 °C were investigated. The emission intensity of the dye at 485 nm increased 100-fold upon the addition of the CD4D12 aggregates in buffer containing free fluorophore without the protein. There was no increase in intensity for thioflavin T when incubated with solutions of CD4D1 (see Figure 2B in the Supporting Information). Electron microscopy confirmed the presence of amyloid-like fibrils for CD4D12 aggregates at pH 7.4. These fibrils had a mean diameter of 20−30 nm and a mean length of 700−950 nm. Most proteins that form amyloids do so at extremes of pH/temperature and/or over a period of several weeks (*51*, *52*). Because CD4D12 forms amyloid precipitates over a period of a week at neutral pH and room temperature, it is a convenient model system to study amyloid formation under physiological conditions.

Characterization of TrxCD4PEP1 and CD4PEP1. Insertion into a loop of thioredoxin greatly increased the yield and purity of CD4PEP1. The molecular weight of the fusion

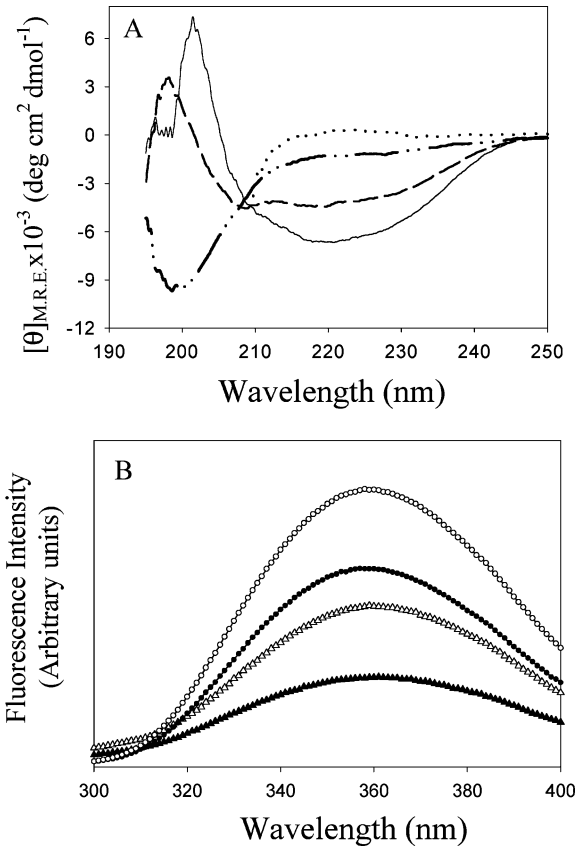

FIGURE 6: Characterization of 21−64 peptide derivatives of CD4D12. (A) Far-UV CD spectra of 50 μM CD4PEP1 in PBS (-···-) or in GdnCl (···), 10 μM thioredoxin (Trx) (—), and TrxCD4PEP1 (- - -). (B) Fluorescence emission spectra for CD4PEP1 (● and ○) and TrxCD4PEP1 (▲ and △) in the native (● and ▲) and unfolded (○ and △) states at pH 7.8. Protein was incubated in native buffer (phosphate buffer at pH 7.8) or 6 M GdnCl for 3 h at 25 °C. Spectra were recorded with excitation at 280 nm.

protein was confirmed using mass spectrometry (expected mass of 19 436.8 Da and observed mass of 19 436.14 Da). The gel-filtration studies show that the protein is monomeric (data not shown). Far-UV CD showed (Figure 6A, - - -) that the fusion protein had a well-defined secondary structure. However, there was some alteration in the secondary structural content of N-HisTrx after insertion of the peptide. To confirm the activity of the fusion protein, both enzyme-linked immunosorbent assay (ELISA) and surface plasmon resonance were performed, which showed that the fusion was able to interact with gp120. N-HisTrx did not show any binding to gp120. To increase the yield further and to characterize CD4PEP1 in the absence of thioredoxin, CD4PEP1 was cloned and expressed in the pMMHa vector as described in the Materials and Methods. Mass spectrometry analysis confirmed the molecular weight of the purified peptide (expected molecular weight of 4883.4 Da and observed molecular weight of 4883.41 Da). The monomeric state of the native peptide was confirmed by gel filtration. The CD spectrum (Figure 6A, -···-) of CD4PEP1 under native conditions shows that it is not folded into a well-defined β sheet. However, there was an appreciable decrease in the magnitude of the MRE when the CD spectrum was recorded for the denatured peptide (Figure 6A, ···) in the presence of 6 M GdnCl, which shows that the peptide has some residual secondary structure in aqueous solution. The tryptic digestion

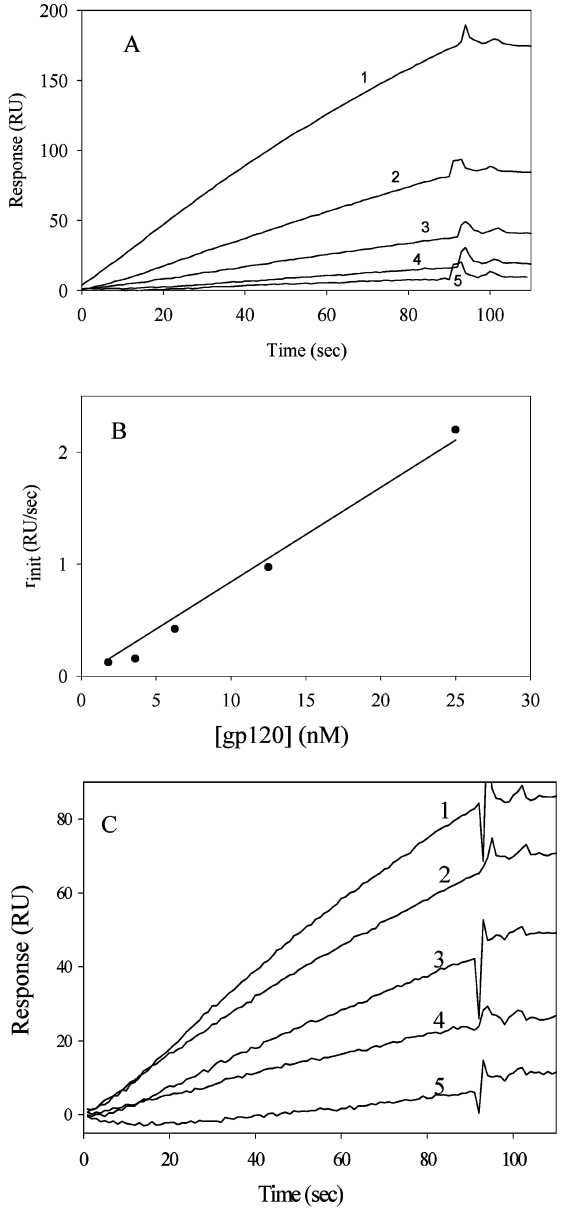

FIGURE 7: (A) Sensorgram overlays for the binding of varying concentrations of gp120 to surface-immobilized F105. Surface density, 1000 RU; buffer, 10 mM phosphate (pH 7.4), 150 mM NaCl, and 0.01% Tween20; flow rate, 30 μL/min. Different concentrations of gp120 [25 nM (1), 12.5 nM (2), 6.1 nM (3), 3.6 nM(4), and 1.8 nM (5)] were passed over the F105 surface. (B) First 30 s of the association phase was fit to a straight line to obtain the initial rate of binding (r_{init}). As expected, r_{init} depends linearly upon the concentration of gp120 and the straight line represents the best fit of r_{init} as a function of [gp120]. The slope of this line can be used to calculate the concentration of gp120 from a measured value of r_{init} as described in the text. (C) Sensorgram overlays for the binding of gp120 preincubated with varying concentrations of CD4D12 to surface-immobilized F105. Surface density, 1000 RU; buffer, 10 mM phosphate (pH 7.4), 150 mM NaCl, and 0.01% Tween20; flow rate, 30 μL/min. A total of 12.5 nM gp120 was incubated alone (1) or with 37 nM (2), 100 nM (3), 200 nM (4), and 400 nM (5) CD4D12 for 30 min before passing over the F105 surface.

pattern of the peptide also confirms the presence of some residual structure (see the later section).

Fluorescence emission spectra of these derivatives were obtained in PBS with and without 6 M GdnCl (Figure 6B). As with CD4PEP1 and CD4D1, there was an increase in intensity of fluorescence in 6 M GdnCl but no red shift, which indicates that, in the absence of the denaturant, the Trp residues in a small fraction of CD4PEP1 molecules are in a structured conformation where quenching by some nearby residue occurs.

Table 1: Dissociation Constants (K_D Values) for the Interaction of gp120 with Various CD4 Analogues

CD4 analogue	apparent K_D (nM)	
	present work	previous studies
rsCD4	12 ± 5	5 (*56*)
CD4D12	15 ± 6	20 (*53*)
CD4D1	40 ± 10	
TrxCD4PEP1	31 ± 4 ($\times 10^3$)	
CD4PEP1	26 ± 3 ($\times 10^3$)	

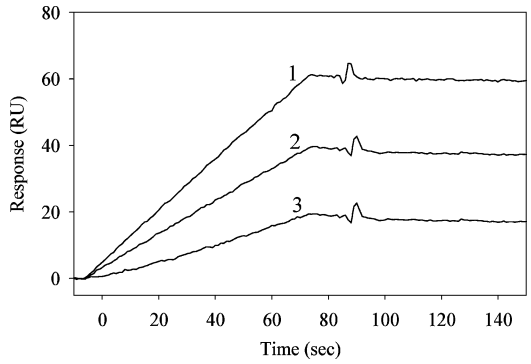

FIGURE 8: Sensorgram overlays for the binding of gp120 preincubated with CD4D12 and CD4 derivatives to surface-immobilized 17b. Surface density, 1000 RU; buffer, 10 mM phosphate (pH 7.4), 150 mM NaCl, and 0.01% Tween20; flow rate, 30 μL/min. A total of 50 nM gp120 was incubated with 250 nM CD4D12 (1), 250 nM CD4D1 (2), and 5 μM CD4PEP1 (3). gp120 did not exhibit significant binding to 17b when incubated with 250 nM CD4PEP1.

Biacore Experiments. K_D values for gp120 binding to CD4D12 as well as two synthetic miniprotein mimics of CD4D12, M9 and M33, have been measured previously by isothermal titration calorimetry for (CD4D12 and M33) (*53*, *54*) and competition ELISA/Biacore for M9 (*32*, *33*). The affinity of gp120 for CD4 analogues in the present study could, in principle, be measured by SPR after immobilization of either binding partner on a chip surface. Unfortunately, this was not possible because gp120 as well as the CD4 analogues became rapidly inactivated upon immobilization. Hence, an alternate strategy using a solution-phase competition binding assay was employed (Figure 7). The results are summarized in Table 1. K_D values for M9 and M33 measured using the competition binding assay were similar to previously published values and are listed in Table 1 of the Supporting Information. The data indicate that CD4D1 binds approximately 3-fold weaker than CD4D12. Although CD4PEP1 binds gp120 relatively weakly, the affinity is approximately 10-fold tighter than CD4-derived peptide fragments studied previously (*55*). CD4PEP1 is also able to trigger the appropriate conformational change in gp120 upon binding as evident from the 17b binding assay (Figure 8). The low affinity of CD4PEP1 is also less likely to be a problem when it is used in a single chain context. In addition, because CD4PEP1 is composed of naturally occurring amino acids, the affinity can be potentially improved using phage display.

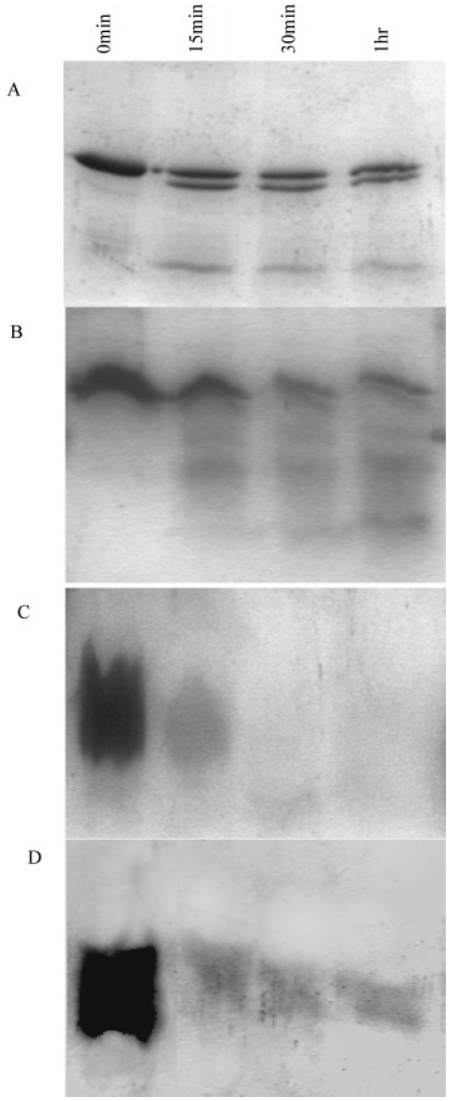

FIGURE 9: SDS−PAGE of proteolytic digests of CD4D12, CD4D1, and CD4PEP1 (at pH 7.8) by trypsin. Proteolysis was done at 37 °C. Samples were taken from the reaction mixtures at the indicated time and mixed with 0.1% formic acid to stop the proteolysis. Samples were boiled with 2% SDS and 100 mM DTT prior to loading. The gel was stained with Coomassie blue. (A) CD4D12, (B) CD4D1, (C) CD4PEP1, and (D) CD4PEP1 in the presence of 6 M sarcosine.

Effect of Osmolyte on the Stability and Structure of CD4D1 and CD4PEP1. Steady-state fluorescence measurements were carried out with CD4D12, CD4D1, and CD4PEP1 in the presence and absence of sarcosine. In the presence of 6 M sarcosine, there was a significant blue shift from 360 to 343 nm for CD4D1 (see curves 1 and 2 in Figure 3 of the Supporting Information). There was no wavelength shift in the case of CD4D12 in the presence and absence of sarcosine (data not shown). The presence of sarcosine resulted in the quenching of fluoresence spectra as shown by the decrease in intensity of emission. The blue shift indicates that the tryptophans contributing to this emission band are less exposed than in the absence of sarcosine, suggesting an increased amount of structure formation for CD4D1 in the presence of sarcosine.

Proteolytic digestion was carried out to probe the protein conformation by looking at the sensitivity of the protein to trypsin cleavage in the presence and absence of sarcosine. Figure 9 shows the tryptic cleavage pattern for CD4D12. After 15 min, three major bands appeared (Figure 9A). One of the bands was approximately the same size as that of the full-length protein. The molecular weight of the second band indicates that cleavage may be either near the N or C terminus of the protein. There was no further change in the cleavage pattern even after 1 h of trypsin digestion. An appreciable fraction of CD4D1 was also not digested (Figure 9B) even after 1 h, confirming that CD4D1 has an appreciable structure. However, no additional protection from trypsin digestion in the presence of 6 M sarcosine could be observed.

CD4PEP1 showed a blue shift in the fluorescence emission spectra from 360 to 345 nm, indicating folding of the peptide in the presence of sarcosine (see curves 3 and 4 in Figure 3 of the Supporting Information). For CD4PEP1, there was some protection against trypsin cleavage in the presence of 6 M sarcosine. Even after 1 h, CD4PEP1 was not completely digested in the presence of sarcosine; however, in the absence of sarcosine, 30 min was sufficient to completely digest the peptide (parts C and D of Figure 9). TrxCD4PEP1 was completely digested within 15 min of incubation with trypsin (data not shown).

CONCLUSIONS

Several derivatives of domain D1 of human CD4 have been constructed and characterized. Mutations were introduced into a fragment consisting of residues 1−99 of human CD4, to minimize aggregation of the protein at exposed hydrophobic sites. The resulting protein (CD4D1) was shown to be monomeric and bound gp120 with an affinity approximately 3-fold weaker than CD4D12. In contrast, a fragment consisting of residues 1−99 without the additional mutations could not be refolded (data not shown). The small reduction in gp120 affinity for CD4D1 relative to CD4D12 is probably because the entire population of CD4D1 is not completely folded. Future studies will explore effects of various engineered disulfides and other stabilizing interactions on the stability and binding properties of CD4D1. A shorter peptide consisting of residues 21−64 of CD4 was also expressed in high yield and characterized. This fragment is less structured than CD4D1 but still retains some residual structure and binds gp120 about 1000-fold weaker than CD4D12. Future studies will attempt to increase the stability and gp120-binding affinity of the fragment. CD4D1 is a useful reagent because it can be expressed to 5−10-fold higher levels than CD4D12 and unlike CD4D12 does not exhibit any tendency to form any amyloid precipitates. Finally, the design strategies outlined in the present work can be used to express single domains of other multidomain proteins.

ACKNOWLEDGMENT

We thank Dr. S. S. Indi for acquiring the E.M. image.

SUPPORTING INFORMATION AVAILABLE

Dissociation constants (K_D values) for the interaction of gp120 with various CD4 analogues (Table 1), ANS-binding studies with native CD4D12 and CD4D1 (Figure 1), CD4D12 aggregates to form amyloid-like fibrils (Figure 2), and effect of osmolyte on the structure of CD4 derivatives in PBS at

pH 7.4 (Figure 3). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Littman, D. R. (1987) The structure of the CD4 and CD8 genes, *Annu. Rev. Immunol. 5*, 561−584.
2. Doyle, C., and Strominger, J. L. (1987) Interaction between CD4 and class II MHC molecules mediates cell adhesion, *Nature 330*, 256−259.
3. Kitchen, S. G., Jones, N. R., LaForge, S., Whitmire, J. K., Vu, B. A., Galic, Z., Brooks, D. G., Brown, S. J., Kitchen, C. M., and Zack, J. A. (2004) CD4 on CD8(+) T cells directly enhances effector function and is a target for HIV infection, *Proc. Natl. Acad. Sci. U.S.A. 101*, 8727−8732.
4. Li, Q. J., Dinner, A. R., Qi, S., Irvine, D. J., Huppa, J. B., Davis, M. M., and Chakraborty, A. K. (2004) CD4 enhances T cell sensitivity to antigen by coordinating Lck accumulation at the immunological synapse, *Nat. Immunol. 5*, 791−799.
5. Dalgleish, A. G., Beverley, P. C., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus, *Nature 312*, 763−767.
6. Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T., and Capon, D. J. (1987) Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor, *Cell 50*, 975−985.
7. Wu, H., Kwong, P. D., and Hendrickson, W. A. (1997) Dimeric association and segmental variability in the structure of human CD4, *Nature 387*, 527−530.
8. Ryu, S. E., Kwong, P. D., Truneh, A., Porter, T. G., Arthos, J., Rosenberg, M., Dai, X. P., Xuong, N. H., Axel, R., Sweet, R. W., et al. (1990) Crystal structure of an HIV-binding recombinant fragment of human CD4, *Nature 348*, 419−426.
9. Wang, J. H., Yan, Y. W., Garrett, T. P., Liu, J. H., Rodgers, D. W., Garlick, R. L., Tarr, G. E., Husain, Y., Reinherz, E. L., and Harrison, S. C. (1990) Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains, *Nature 348*, 411−418.
10. Ryu, S. E., Truneh, A., Sweet, R. W., and Hendrickson, W. A. (1994) Structures of an HIV and MHC binding fragment from human CD4 as refined in two crystal lattices, *Structure 2*, 59−74.
11. Wang, J. H., Meijers, R., Xiong, Y., Liu, J. H., Sakihama, T., Zhang, R., Joachimiak, A., and Reinherz, E. L. (2001) Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule, *Proc. Natl. Acad. Sci. U.S.A. 98*, 10799−10804.
12. Kwong, P. D., Wyatt, R., Majeed, S., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (2000) Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates, *Struct. Fold Des. 8*, 1329−1339.
13. Hoffman, T. L., LaBranche, C. C., Zhang, W., Canziani, G., Robinson, J., Chaiken, I., Hoxie, J. A., and Doms, R. W. (1999) Stable exposure of the coreceptor-binding site in a CD4-independent HIV-1 envelope protein, *Proc. Natl. Acad. Sci. U.S.A. 96*, 6359−6364.
14. Berger, E. A., Fuerst, T. R., and Moss, B. (1988) A soluble recombinant polypeptide comprising the amino-terminal half of the extracellular region of the CD4 molecule contains an active binding site for human immunodeficiency virus, *Proc. Natl. Acad. Sci. U.S.A. 85*, 2357−2361.
15. Siliciano, R. F., Lawton, T., Knall, C., Karr, R. W., Berman, P., Gregory, T., and Reinherz, E. L. (1988) Analysis of host-virus interactions in AIDS with anti-gp120 T cell clones: Effect of HIV sequence variation and a mechanism for CD4+ cell depletion, *Cell 54*, 561−575.
16. Smith, D. H., Byrn, R. A., Marsters, S. A., Gregory, T., Groopman, J. E., and Capon, D. J. (1987) Blocking of HIV-1 infectivity by a soluble, secreted form of the CD4 antigen, *Science 238*, 1704−1707.
17. Davis, C. B., Boyle, K. E., Urbanski, J. J., Paradysz, R. T., and Fong, K. L. (1992) Disposition of metabolically labeled recombinant soluble CD4 (sT4) in male Sprague−Dawley rats following intravenous and subcutaneous administration, *Drug Metab. Dispos. 20*, 695−705.
18. Ashkenazi, A., Smith, D. H., Marsters, S. A., Riddle, L., Gregory, T. J., Ho, D. D., and Capon, D. J. (1991) Resistance of primary isolates of human immunodeficiency virus type 1 to soluble CD4 is independent of CD4−rgp120 binding affinity, *Proc. Natl. Acad. Sci. U.S.A. 88*, 7056−7060.
19. Clapham, P. R., McKnight, A., and Weiss, R. A. (1992) Human immunodeficiency virus type 2 infection and fusion of CD4-negative human cell lines: Induction and enhancement by soluble CD4, *J. Virol. 66*, 3531−3537.
20. Fouts, T., Godfrey, K., Bobb, K., Montefiori, D., Hanson, C. V., Kalyanaraman, V. S., DeVico, A., and Pal, R. (2002) Cross-linked HIV-1 envelope-CD4 receptor complexes elicit broadly cross-reactive neutralizing antibodies in rhesus macaques, *Proc. Natl. Acad. Sci. U.S.A. 99*, 11842−11847. Epub 12002 Aug 11821.
21. Varadarajan, R., Sharma, D., Chakraborty, K., Patel, M., Citron, M., Sinha, P., Yadav, R., Rashid, U., Kennedy, S., Eckert, D., Geleziunas, R., Bramhill, D., Schleif, W., Liang, X., and Shiver, J. (2005) Characterization of gp120 and its single-chain derivatives, gp120-CD4D12 and gp120-M9: Implications for targeting the CD4i epitope in human immunodeficiency virus vaccine design, *J. Virol. 79*, 1713−1723.
22. Chang, T. L., Chang, C. H., Simpson, D. A., Xu, Q., Martin, P. K., Lagenaur, L. A., Schoolnik, G. K., Ho, D. D., Hillier, S. L., Holodniy, M., Lewicki, J. A., and Lee, P. P. (2003) Inhibition of HIV infectivity by a natural human isolate of *Lactobacillus jensenii* engineered to express functional two-domain CD4, *Proc. Natl. Acad. Sci. U.S.A. 100*, 11672−11677.
23. Jager, M., and Pluckthun, A. (1999) Folding and assembly of an antibody Fv fragment, a heterodimer stabilized by antigen, *J. Mol. Biol. 285*, 2005−2019.
24. Traunecker, A., Dolder, B., Oliveri, F., and Karjalainen, K. (1989) Solubilizing the T-cell receptor—Problems in solution, *Immunol. Today 10*, 29−32.
25. Traunecker, A., Schneider, J., Kiefer, H., and Karjalainen, K. (1989) Highly efficient neutralization of HIV with recombinant CD4-immunoglobulin molecules, *Nature 339*, 68−70.
26. Tendian, S. W., Myszka, D. G., Sweet, R. W., Chaiken, I. M., and Brouillette, C. G. (1995) Interdomain communication of T-cell CD4 studied by absorbance and fluorescence difference spectroscopy measurements of urea-induced unfolding, *Biochemistry 34*, 6464−6474.
27. Moebius, U., Pallai, P., Harrison, S. C., and Reinherz, E. L. (1993) Delineation of an extended surface contact area on human CD4 involved in class II major histocompatibility complex binding, *Proc. Natl. Acad. Sci. U.S.A. 90*, 8259−8263.
28. Dianzani, U., Shaw, A., al-Ramadi, B. K., Kubo, R. T., and Janeway, C. A., Jr. (1992) Physical association of CD4 with the T cell receptor, *J. Immunol. 148*, 678−688.
29. Rudd, C. E., Trevillyan, J. M., Dasgupta, J. D., Wong, L. L., and Schlossman, S. F. (1988) The CD4 receptor is complexed in detergent lysates to a protein-tyrosine kinase (pp58) from human T lymphocytes, *Proc. Natl. Acad. Sci. U.S.A. 85*, 5190−5194.
30. Staley, J. P., and Kim, P. S. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor, *Protein Sci. 3*, 1822−1832.
31. 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.
32. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Menez, A., Roumestand, C., Yang, Y. S., Ylisastigui, L., Benjouad, A., and Gluckman, J. C. (1999) Rational engineering of a miniprotein that reproduces the core of the CD4 site interacting with HIV-1 envelope glycoprotein, *Proc. Natl. Acad. Sci. U.S.A. 96*, 13091−13096.
33. Martin, L., Stricher, F., Misse, D., Sironi, F., Pugniere, M., Barthe, P., Prado-Gotor, R., Freulon, I., Magne, X., Roumestand, C., Menez, A., Lusso, P., Veas, F., and Vita, C. (2003) Rational design of a CD4 mimic that inhibits HIV-1 entry and exposes cryptic neutralization epitopes, *Nat. Biotechnol. 21*, 71−76.
34. Posner, M. R., Cavacini, L. A., Emes, C. L., Power, J., and Byrn, R. (1993) Neutralization of HIV-1 by F105, a human monoclonal antibody to the CD4 binding site of gp120, *J. Acquired Immune Defic. Syndr. 6*, 7−14.
35. Chakraborty, K., Shivakumar, P., Raghothama, S., and Varadarajan, R. (2005) NMR structural analysis of a peptide mimic of the bridging sheet of HIV-1 gp120 in methanol and water, *Biochem. J. 390*, 573−581.

36. Ratnaparkhi, G. S., and Varadarajan, R. (2001) Osmolytes stabilize ribonuclease S by stabilizing its fragments S protein and S peptide to compact folding-competent states, *J. Biol. Chem. 276*, 28789−28798.
37. Mizukami, T., Fuerst, T. R., Berger, E. A., and Moss, B. (1988) Binding region for human immunodeficiency virus (HIV) and epitopes for HIV-blocking monoclonal antibodies of the CD4 molecule defined by site-directed mutagenesis, *Proc. Natl. Acad. Sci. U.S.A. 85*, 9273−9277.
38. Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody, *Nature 393*, 648−659.
39. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP: A structural classification of proteins database for the investigation of sequences and structures, *J. Mol. Biol. 247*, 536−540.
40. Chao, B. H., Costopoulos, D. S., Curiel, T., Bertonis, J. M., Chisholm, P., Williams, C., Schooley, R. T., Rosa, J. J., Fisher, R. A., and Maraganore, J. M. (1989) A 113-amino acid fragment of CD4 produced in *Escherichia coli* blocks human immunodeficiency virus-induced cell fusion, *J. Biol. Chem. 264*, 5812−5817.
41. Moebius, U., Clayton, L. K., Abraham, S., Harrison, S. C., and Reinherz, E. L. (1992) The human immunodeficiency virus gp120 binding site on CD4: Delineation by quantitative equilibrium and kinetic binding studies of mutants in conjunction with a high-resolution CD4 atomic structure, *J. Exp. Med. 176*, 507−517.
42. Abdulaev, N. G., Ngo, T., Chen, R., Lu, Z., and Ridge, K. D. (2000) Functionally discrete mimics of light-activated rhodopsin identified through expression of soluble cytoplasmic domains, *J. Biol. Chem. 275*, 39354−39363.
43. Tripp, B. C., Lu, Z., Bourque, K., Sookdeo, H., and McCoy, J. M. (2001) Investigation of the "switch-epitope" concept with random peptide libraries displayed as thioredoxin loop fusions, *Protein Eng. 14*, 367−377.
44. Chakrabarti, A., Srivastava, S., Swaminathan, C. P., Surolia, A., and Varadarajan, R. (1999) Thermodynamics of replacing an α-helical Pro residue in the P40S mutant of *Escherichia coli* thioredoxin, *Protein Sci. 8*, 2455−2459.
45. Clark, P. L., Liu, Z. P., Zhang, J., and Gierasch, L. M. (1996) Intrinsic tryptophans of CRABPI as probes of structure and folding, *Protein Sci. 5*, 1108−1117.
46. Wu, J., Yang, J. T., and Wu, C. S. (1992) β-II conformation of all-β proteins can be distinguished from unordered form by circular dichroism, *Anal. Biochem. 200*, 359−364.
47. Sanchez-Puig, N., Veprintsev, D. B., and Fersht, A. R. (2005) Human full-length Securin is a natively unfolded protein, *Protein Sci. 14*, 1410−1418.
48. Denning, D. P., Uversky, V., Patel, S. S., Fink, A. L., and Rexach, M. (2002) The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein, *J. Biol. Chem. 277*, 33447−33455. Epub 32002 Jun 33413.
49. Inouye, H., and Kirschner, D. A. (2000) A β fibrillogenesis: Kinetic parameters for fibril formation from congo red binding, *J. Struct. Biol. 130*, 123−129.
50. Tjernberg, L. O., Callaway, D. J., Tjernberg, A., Hahne, S., Lilliehook, C., Terenius, L., Thyberg, J., and Nordstedt, C. (1999) A molecular model of Alzheimer amyloid β-peptide fibril formation, *J. Biol. Chem. 274*, 12619−12625.
51. Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G., and Dobson, C. M. (1999) Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils, *Proc. Natl. Acad. Sci. U.S.A. 96*, 3590−3594.
52. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid fibril formation by an SH3 domain, *Proc. Natl. Acad. Sci. U.S.A. 95*, 4224−4228.
53. Myszka, D. G., Sweet, R. W., Hensley, P., Brigham-Burke, M., Kwong, P. D., Hendrickson, W. A., Wyatt, R., Sodroski, J., and Doyle, M. L. (2000) Energetics of the HIV gp120−CD4 binding reaction, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9026−9031.
54. Huang, C. C., Stricher, F., Martin, L., Decker, J. M., Majeed, S., Barthe, P., Hendrickson, W. A., Robinson, J., Roumestand, C., Sodroski, J., Wyatt, R., Shaw, G. M., Vita, C., and Kwong, P. D. (2005) Scorpion-toxin mimics of CD4 in complex with human immunodeficiency virus gp120 crystal structures, molecular mimicry, and neutralization breadth, *Structure 13*, 755−768.
55. Gizachew, D., Moffett, D. B., Busse, S. C., Westler, W. M., Dratz, E. A., and Teintze, M. (1998) NMR studies on the conformation of the CD4 36−59 peptide bound to HIV-1 gp120, *Biochemistry 37*, 10616−10625.
56. Zhang, W., Canziani, G., Plugariu, C., Wyatt, R., Sodroski, J., Sweet, R., Kwong, P., Hendrickson, W., and Chaiken, I. (1999) Conformational changes of gp120 in epitopes near the CCR5 binding site are induced by CD4 and a CD4 miniprotein mimetic, *Biochemistry 38*, 9405−9416.

BI051120S