A Common Pharmacophoric Footprint for AIDS Vaccine Design

Christoph Pisterer,[†] Dan Mihailescu,[†,‡] Jeremy C. Smith,[†] and Jennifer Reed*[,‡]

Computational Molecular Biophysics, Interdisciplinary Center for Computational Science (IWR), University of Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany, and German Cancer Research Center (Deutsches Krebsforschungszentrum), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Received December 4, 2003

The most promising target antigen for an HIV vaccine designed using the classic antibody strategy has been the viral coat protein gp120. Unfortunately, its high variability has prevented this approach. We examine here a 15-residue peptide derived from the CD4-binding domain of gp120. By use of molecular dynamics computer simulation, it is shown that despite considerable sequence variation, the three-dimensional structure of the peptide is preserved over the full range of clade-specific sequences. Furthermore, sequences threaded onto the structure exhibit common three-dimensional electrostatic and hydrophobic properties. These common physicochemical characteristics constitute a pharmacophoric footprint that promises to be useful in the design of a synthetic antigen for vaccine development.

Introduction

The attitude toward the utility of vaccine development in combating the AIDS epidemic has come full circle since 1984. Initial attempts to halt the spread of the disease were all vaccine-based, understandably since this approach had been consistently successful in the past.[1,2] The hope, as with other viruses, was to induce neutralizing antibodies that would prevent entry into the host's target cells. It rapidly became clear, however, that the conventional antigen (the viral coat protein) displayed a level of variability beyond that of any virus hitherto encountered[3,4] and that the most antigenic epitopes of the coat protein were also the most variable.[5] When antibodies were produced, they did not fully protect and were only temporarily effective; the virus produced insensitive mutations more quickly than the immune system could respond. Employing more conserved viral proteins as antigens, such as p24, which is used as a marker of infection, was also unsuccessful because these occur in the serum principally after cell lysis, after the damage is done.

For the above reasons, after about 1990 work on AIDS vaccines tended to lapse in favor of a treatment-oriented approach. The development of reverse transcriptase inhibitors and HIV protease inhibitors and the definite, if limited, success of HAART (highly active antiretroviral therapy) have justified this approach. It has become clear, however, that HAART cannot be applied on the worldwide basis needed to combat AIDS; it is too expensive, it requires rigorous adherence to a complex schedule on the part of the patient, and there are serious side effects associated with long-term use. Recently, development of an AIDS vaccine has again become a priority, armed with our present deeper appreciation of the difficulties involved. Further, the recent concentration on stimulating cellular immunity is being relaxed while the classic antibody strategy is receiving renewed attention.[6]

The ideal antigen for a vaccine designed using the classic antibody strategy remains the viral coat protein gp120. The chief obstacles that gp120 presents as an antigen are its high variability and the questionable accessibility to antibodies of those epitopes that might be expected to be functionally conserved, i.e., those at the binding site for the T-cell receptor CD4. However, the partial cross-reactivity of an antibody to a binding site epitope, b12,[7] indicates that targeting a functionally conserved epitope can be an effective means of avoiding the variability problem as well as demonstrating that at least part of the CD4-binding site is exposed even in the trimeric configuration of gp120 found in the virus. Thus, minimally variable or nonvariable epitopes corresponding to or overlapping the area targeted by the b12 antibody are a promising lead for vaccine development.

The questions addressed in the present work are whether a 15-residue peptide derived from the CD4-binding domain of gp120, partially overlapping the b12 epitope and independently capable of specific binding to the receptor, is structured in aqueous solution; whether this structure is preserved over the full range of clade-specific sequences; and if so, whether the common structure also presents common physicochemical characteristics that might be useful in the design of a synthetic antigen provoking a cross-clade response. (Although point mutations at sites distant from this sequence can affect binding affinity, suggesting that in the native fold the region in contact with the CD4 receptor includes some side chains external to the sequence, these 15 residues are the only gp120 fragment capable of independent binding to the receptor in isolation and thus constitute the major binding domain.) Our previous experimental work on the peptide indicated that the secondary structure in this region was conserved despite significant sequence variability.[8] If this backbone structure can be determined and the full

* To whom correspondence should be addressed. Phone: 0049-6221-423256. Fax: 0049-6221-423259. E-mail: j.reed@dkfz-heidelberg.de.
[†] University of Heidelberg.
[‡] German Cancer Research Center.

10.1021/jm031125s CCC: $27.50
Published on Web 06/18/2004

range of clade-specific sequences can be threaded onto it, then examination of the average molecular surface presented may reveal those conserved physicochemical characteristics required for binding to CD4. These would in turn provide a template for the design of an antigen that may provoke a broad or even fully cross-reactive response.

The peptide of interest is not suitable for experimental structure determination using NMR spectroscopy in aqueous solution. However, an alternative method exists, molecular dynamics (MD) computer simulation, which is now sufficiently developed to produce reliable information on solution peptide structures.[9] Computer simulations played an important role in the development of HIV-protease and reverse transcriptase inhibitors for HAART.[10]

We have previously reported an MD study of a single sequence that demonstrated that the sequence adapts a nonrandom structure in aqueous solution[11] and have characterized the structure in detail.[12] Here, after subjecting that model to a series of tests controlling its uniqueness and stability over the full range of clade sequences, we present the resulting electrostatic and hydrophobic surfaces for each clade and discuss the features defining the pharmacophoric footprint conserved within the variability. These features can now be used in the design of a strain-independent antigen.

Results

The antigen-combining site of an antibody is complementary in shape to the three-dimensional surface of the corresponding epitope. Binding is also determined by the complementarity of the molecular electrostatic potential at the surface and by interactions between hydrophobic residues. Consequently, to determine whether a physicochemical pharmacophoric footprint might exist in the gp120 segment under study that could form a basis for vaccine design, the three-dimensional structure, if any, of the peptide must be found and the molecular surface characterized in terms of any common electrostatic or hydrophobic characteristics.

Three-Dimensional Structure. The three-dimensional structure of the LAV-strain sequence LPCRIKQFINMWQEV was obtained previously from MD simulations.[11,12] Three simulations were performed in explicit solvent using starting configurations determined by nuclear magnetic resonance (NMR) spectroscopy in three different organic solvents. The three simulations converged to a common structure in both Cartesian and in backbone dihedral conformational space. In Figure 1 is shown the average converged structure from these simulations.

Surface Electrostatic Characteristics: Individual and Conserved. Nineteen different representative sequences from eight different HIV clades (highlighted in Table 1) were threaded onto the molecular structure shown in Figure 1 (see Experimental Section). Figure 2 shows the electrostatic surfaces of these peptides at the 0° face, the face opposite the points of attachment to the main protein and most exposed to the medium. In general, the electrostatic potential surface is strikingly variable. It is clear that an immune response mounted against this epitope in strain no. 4, for example, would have little cross reactivity with strain no. 1. Yet invariant physicochemical features must be present in these peptides to allow recognition by and binding to the CD4 receptor. To reveal these features, the constant signal beneath the variant noise must be determined. To do this, the average solvent-accessible surface and the average electrostatic field were computed. As in classic signal averaging, variably charged areas will tend to cancel out while those that persist throughout the strains will be positively reenforced. The results are presented in Figure 3 for all six orthogonal faces of the average structure. Most of

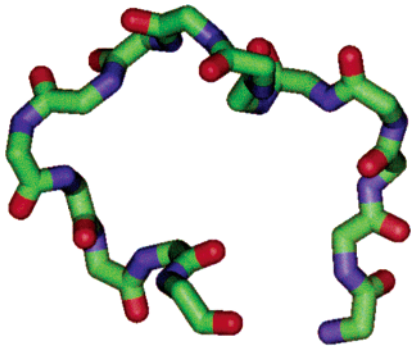

Figure 1. Backbone structure of the 15-residue CD4-binding region from strain LAV in aqueous solution, derived from molecular dynamics simulation.

Table 1. HIV Clade Sequences

sequence	GenBank No.
	Group M
	Subtype A
LQCRIKQIINMWQTV	U56274−U56283
	Subtype B
LPCRIKQIVNMWQEV	U56263−U56268, U56271
LPCRIKQIINMWQEV	U562272, U31364
LPCKIKQIVNMWQEV	U56269, U56270
LPCRIKQTINTWQEV	U562273
LQCRIKQIINMWQEV	U56288, U56290, U56291, U56317−U563321
LQCRIKPIINMWQEV	U56289
LPCRIKQIVNMWQEV	U56299
	Subtype C
IPCRIKQIINMWQEV	U29179, U29696, U31363
IPCRIKQIVNLWQEV	U29694
IPCRIKQIINMWQQV	U29697
IPCRIKQIINMWPEV	U29698
IPCRIKQIINMWQGV	U31362
	Subtype D
LPCRIKQIINMWQEV	U56304, U56311
LPCRIKPIINMWQEV	U56310, U56312
	Subtype E
LPCRIKQIINMWQGV	U56309, U56329, U48272
LPCKIKQIVNMWQGA	U48273
LPCKIKQIINMWQGA	U482274−U48278, U48265, U48267
LPCKIKQIIKMWQGV	U48264, U48269
LPCKIKQIINMWQGV	U48266, U48268
	Subtype F
LLCKIKQIVNLWQEV	U37032
LPCKIRQIVNMWQEV	U37033
	Subtype G
LPCRIKQIVRMWQRV	U08368
LPCKIKQIVRMWQRV	U33095, U08358, U08362
LPCRIKQIIRMWQRV	U08355, U08357, U08359−U08361
LPCRIKQIIIMWQRV	U08363
LPCKIKQIVRMWQRV	U08365
LPCRLKQIVRMWQRV	U08366
LPRRLKQIVRMWQRV	U08367
	Subtype H
LPCRIKQIIRMWQRV	U08355, U08357
LQCRIKQIVNMWQRV	U33104

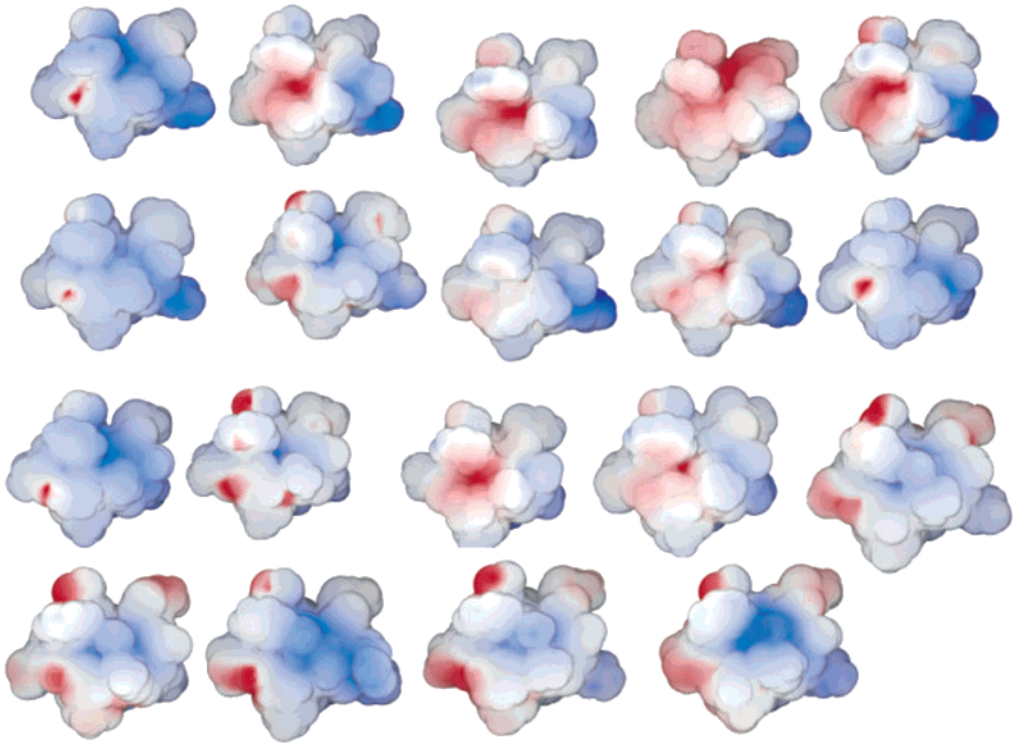

Figure 2. Potential surfaces of 19 strains from 8 clades (see Table 1) threaded onto the backbone structure of Figure 1, seen from the 0° face. An arbitrary set of coordinates was established with 0° corresponding to the face presented externally from the parent protein (gp120) and 90° corresponding roughly to the face connected to the parent protein. Negative charge is shown as red; positive charge is shown as blue. White surfaces are neutral. Scale runs from −4 to +4 V.

the negative potential has been canceled out, although patches of negative potential remain associated with backbone carbonyls. The outstanding feature is an irregular band of positive electrostatic potential spanning $^3/_4$ of the equatorial region of the domain. Thus, the variability of gp120 in the CD4-binding region does in fact mask a conserved pharmacophoric footprint.

How unique are these characteristics to HIV? Would any ensemble of 15-residue sequences threaded onto this structure also show conserved features other than those due to backbone carbonyls and amides? To perform this control, 50 random sequences were threaded onto the average structure and their average electrostatic surface potential was calculated. The surfaces were generated by randomly choosing a number between 1 and 20, each number being associated with a specific amino acid residue. Figure 4 shows the average electrostatic surface of the 50 random sequences, with the amide and carbonyl charges damped to clarify side chain charge distribution. This control computation displays no constant features other than those associated with the backbone.

Surface Hydrophobicity Profiles: Individual and Conserved. Similar calculations were carried out for hydrophobic mapping of the surfaces of the 19 clades, and constant features were revealed through averaging as indicated above. The 0° faces are presented in Figure 5, and the averaged surface is shown from all six faces in Figure 6. Interestingly, the extreme variability seen in the electrostatic surface is not present in the hydrophobic or hydrophilic surface, the strains all showing a strong family resemblance to one another. Unsurprisingly, these features are emphasized in the averaged surface. The "lower" apex is strongly hydrophobic, and the upper apex hydrophilic. The 0° face is dominated by hydrophobic surfaces with the conserved tryptophan residue prominently displayed, while a hydrophilic band around the equatorial region corresponds to the areas of conserved positive potential. This combination of conserved hydrophobic and electrostatic features defines a specific physicochemical constellation, the pharmacophoric footprint.

Cross-Clade Conservation of Backbone Structure Used. The structures given above were derived by assuming, from circular dichroism (CD) studies, that the secondary structure of the backbone was conserved.[8] However, the broadly constant secondary structure seen using CD cannot detect potentially immunologically significant rearrangements of the backbone. In the calculations presented above, a single structure was used for threading—that from the strain LAV. Is this structure also preferred when highly variant sequences are threaded onto it? To check this point, a final control set of calculations was performed in which the 11 most different clade sequences threaded onto the LAV-derived backbone were subjected to MD simulation in explicit solvent (see Experimental Section). Figure 7 shows the superimposed backbones of the N- and C-terminal fragments of the 11 strains at the end of the simulation. The 11 strains clearly adopt a common three-dimensional structure. As was found in the previous work on LAV alone, the N- and C-terminal segments each adopt a preferred structure, with a hinge section in the center of the peptide.[11,12] The LAV backbone from Figure 1 is in black and is mostly not visible in the figure because of its occupation of the same configurational space as the other sequences. Thus, the simulations suggest that the presence of different, naturally occurring side chains

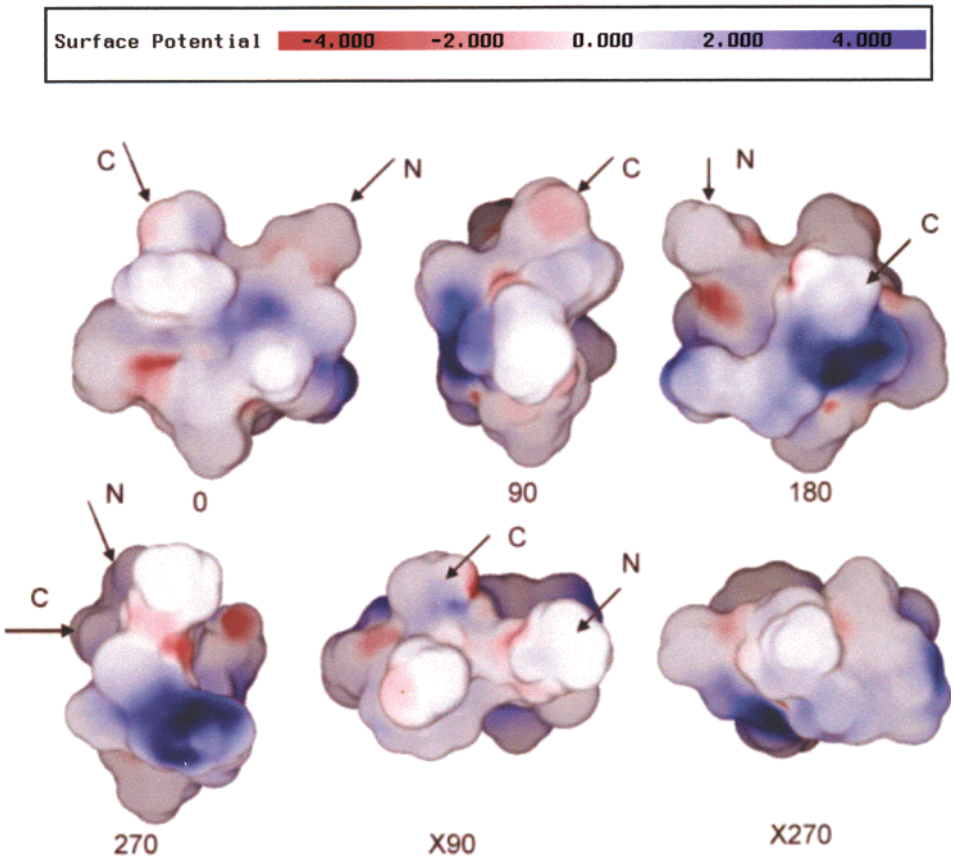

Figure 3. Averaged solvent-accessible and potential surface for the 15-residue CD4-binding region. Color key as in Figure 2. All six orthoganal faces are shown. The first four are ordered following a clockwise rotation around the central vertical axis; ×90 corresponds to the "top" and ×270 corresponds to the "bottom". Black arrows indicate the point of attachment of the N and C termini of the sequence to the parent protein.

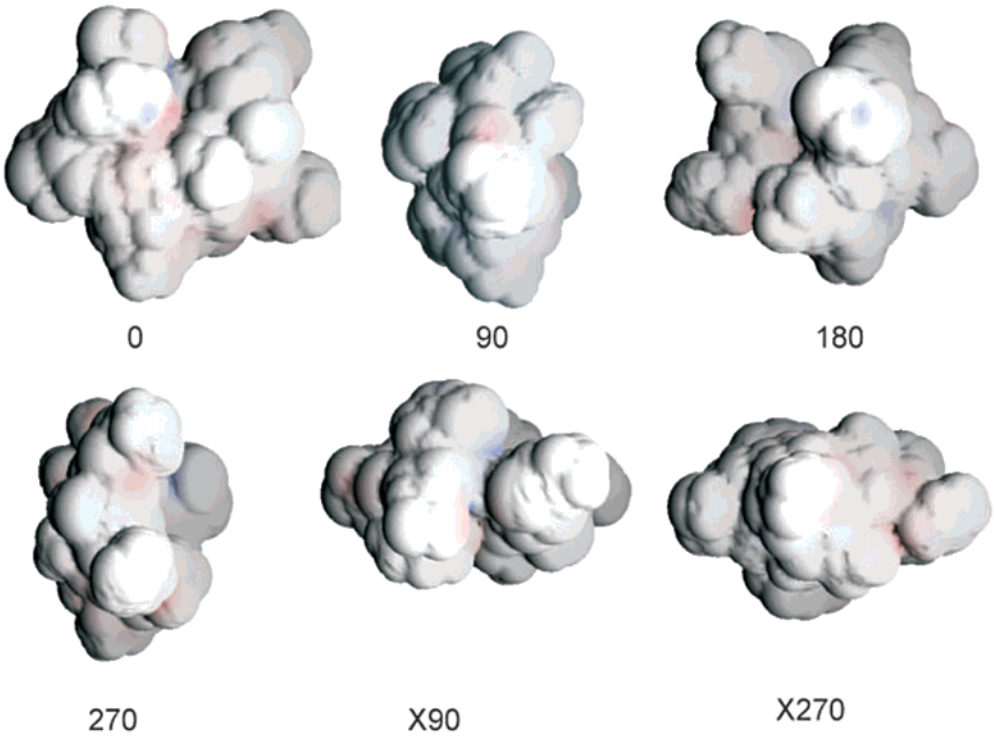

Figure 4. Averaged potential surface for 50 random 15-residue sequences threaded onto the backbone in Figure 1. Backbone carbonyl and amide charges have been suppressed to emphasize any side chain contribution.

does not significantly affect the backbone structure of the peptide. Moreover, the N- and C-terminal residues were not constrained in the simulations, whereas in vivo they are constrained because of their attachment to the

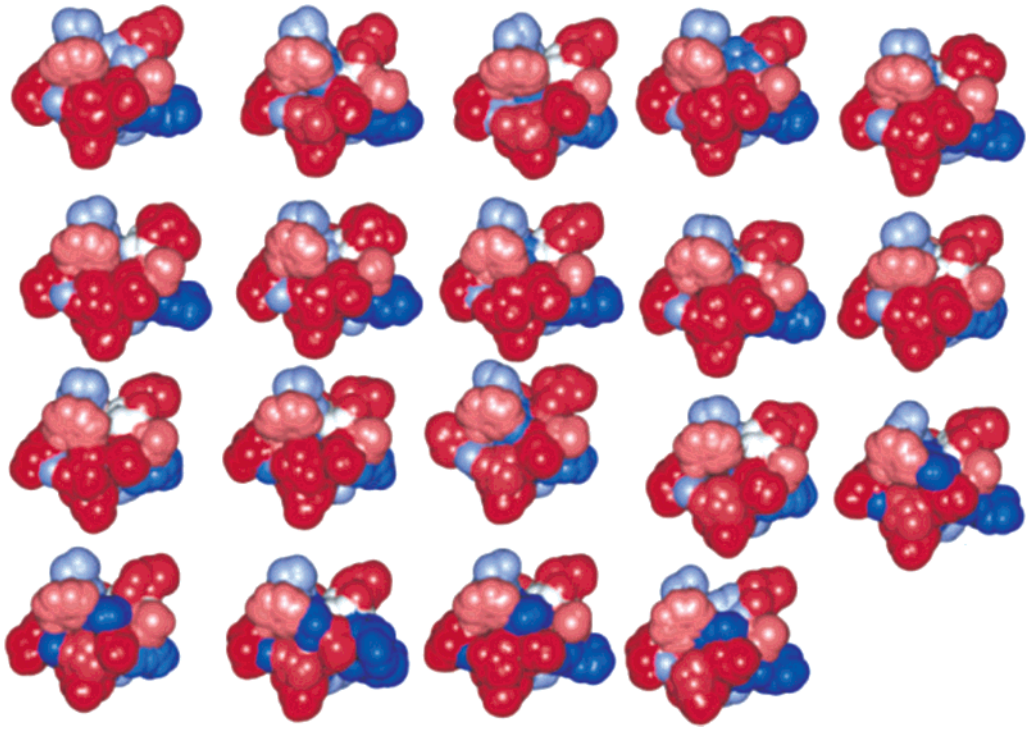

Figure 5. Surface polarity maps of the 19 strains shown in Figure 2. Hydrophobic areas are shown as red; hydrophilic areas are shown as blue.

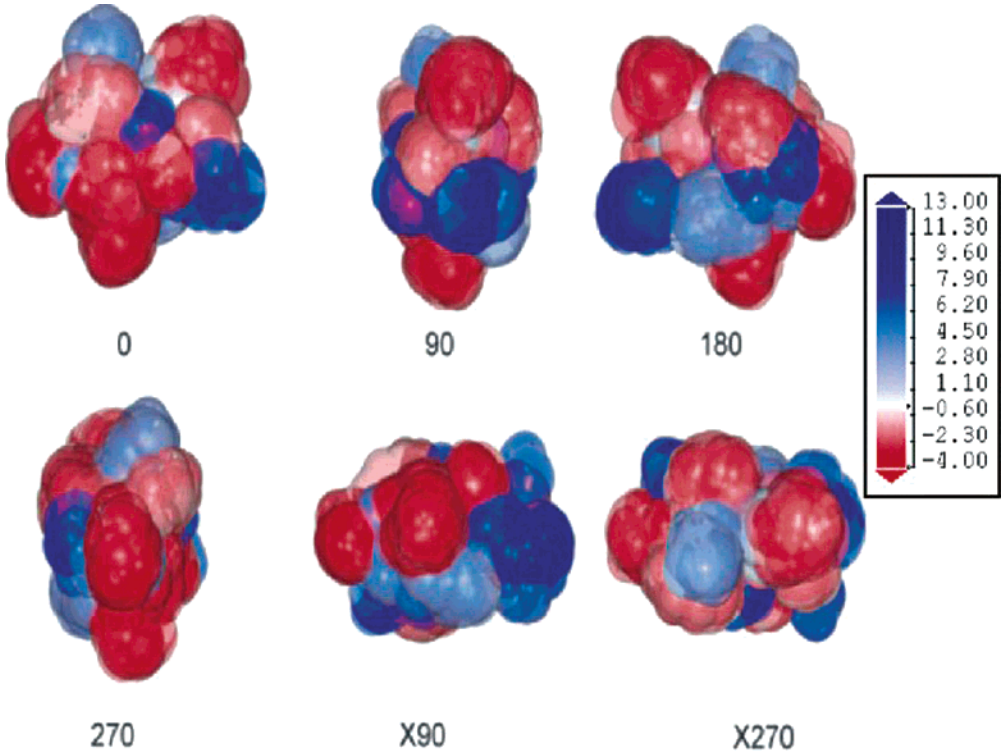

Figure 6. Averaged surface polarity shown from all six orthoganal faces. To generate this figure, the Engelman−Steitz residue hydrophobicity was mapped onto the solvent-accessible surface of each residue and colored red for hydrophobic residues and blue for hydrophilic residues. The 19 surfaces thus calculated (for each sequence) are superposed on the figure using a common backbone. Averaging was performed using a transparency index. Areas of similar hydrophobicity reinforce to display a solid appearance.

parent protein. The use here of the LAV-derived backbone for threading multiple strains thus appears to be justified.

Discussion

Even the most variable of proteins is constrained in this variability when the necessity exists for interacting with a partner molecule. Receptors, binding sites, etc. are designed to recognize a particular pattern of form, surface charge, and surface polarity. Any departure from this will render the protein nonfunctional. Provided the pattern is presented, however, it is largely irrelevant by which chemical compound the pattern was generated. Thus, the taste receptor for sweetness binds

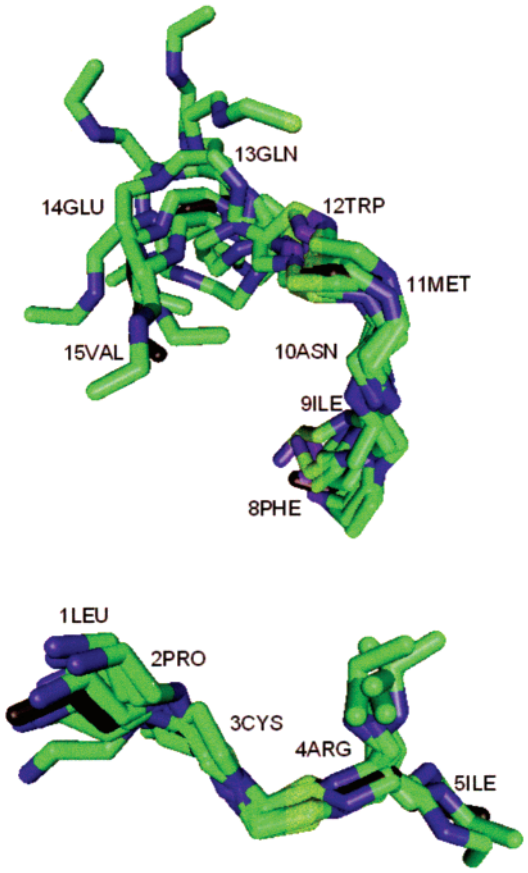

Figure 7. Superimposition of converged N and C termini resulting from MD simulation in explicit solvent of the 11 most divergent sequences. The LAV backbone structure from Figure 1 is shown in black and is mostly not visible because it is virtually identical to the average.

and responds to aspartame as though it were a sugar although it is chemically unrelated.

Despite the notorious variability of the HIV 1 envelope protein gp120, certain portions of it, such as the CD4 binding domain, are constrained in this way. The 15-residue segment at the core of the domain examined here is capable of binding to CD4 on its own.[8] This segment has historically been considered one of the less variable sections of gp120. However, although inspection of this sequence from all known clades reveals a family resemblance, only one amino acid (tryptophan at position 12) is actually retained throughout all viral strains reported to date. Despite this, these strains all bind to the CD4 receptor. A distinct pattern, or pharmacophoric footprint, must then exist that is common to all HIV1 variants.

A clue as to how this might be occurring was that circular dichroism spectra of synthetic peptides incorporating the sequence from different viral strains indicate that the secondary structure in this region is conserved. It has been possible to define this backbone structure using NMR and MD. The calculations involved are standard molecular dynamics simulations of a 15-residue peptide at room temperature. Consequently, it is not guaranteed that all relevant conformations of the peptide are sampled in the accessible simulation time. Yet in previous simulations, the LAV-strain peptide, starting from different experimental geometries, converged toward the same structure, Moreover, although only one of the residues is completely conserved, there is a clear family similarity among the strain sequences. Therefore, it is likely that the present simulated structure constitutes a meaningful conformation of the peptides.

Threading the full spectrum of known clade sequences onto this backbone should, in theory, allow one to see whether a constant physicochemical pattern is present and what it might be. The results of this approach are enlightening with regard to a possible strategy for vaccine design. The difficulty of raising a vaccine against even "constant" segments of gp120 becomes apparent when one looks at the surfaces presented by the CD4 binding sites of different strains given in Figure 2. At first glance, the variability of charge distribution alone makes it appear unlikely that a common pattern is present. This confusion to the eye, which masks the common features seen after averaging (Figure 3), is analogous to the confusion to the immune system, where strong and idiosyncratic charge patches are chosen as favored epitopes while the constant areas are less prominent. The role played by negative side chain charges is especially ingenious. Although they are a conspicuous feature in many strains, they cancel out completely when averaged over the clades. They act as a red herring that lures the immune system down a futile antigenic pathway.

Once stripped of its variant features, the surface of the CD4 binding site displays a highly distinctive charge and polarity distribution. That these are specific to the gp120 sequences and not an artifact of the averaging algorithm used is seen from the featureless surface obtained when the same technique is applied to random 15-mers. The coordinates of the averaged surface with its conserved electrostatic and hydrophobic or hydrophilic characteristics constitute a template for the design of a strain-independent antigen. The 0° and ×270° aspects, as those most exposed in the parent protein, are therefore likely to be the most promising for this purpose.

It should be noted that the entire 15-residue sequence is larger than the size preferred for antigenic epitopes. In principle, the conserved surface features of the N- or C-terminal region alone should be sufficient (Figure 7). One approach would therefore be to raise antibodies against each separately as a combined immunization.

We have attempted to approach the problem of variability in the envelope protein gp120 by using combined biophysical and computational techniques to expose the conserved pharmacophoric footprint underlying the manifold differences in the solvent-exposed surface of the major CD4-binding domain. The coordinates for the averaged electrostatic and hydrophobic surfaces are being made available as an electronic file from the authors. We suggest that they will prove to be useful in the design of a synthetic antigen capable of inducing a broadly cross-reactive immune response to the AIDS virus.

Experimental Section

Selection of Clade Sequences. A sequence search was carried out using the Los Alamos data bank (http://los_alamos/COMPENDIUM/1996/PART-1/summarytables.pdf) for all available sequences of the switch region for which the clade identification has been made. In many cases, sequences in this

region from multiple strains within a clade are identical, although for most clades several different sequences occur. The sequences chosen for this study were those most often reported within a clade and most nearly representative of it, with at least one sequence used for each clade for which sequences had been reported. Nineteen sequences from eight different clades were chosen. Table 1 lists all the different sequences reported for this region, with those used in the present study highlighted.

Computational Modeling Techniques. The three-dimensional structure of the LAV-strain sequence LPCRIKFINMWQEV (Figure 1) was obtained from MD simulations.[11,12] The 19 sequences were threaded onto the backbone structure shown in Figure 1 using the CHARMM software.[13] This involved mutating the residues while keeping all the unchanged atoms fixed, reconstructing the coordinates of the remaining atoms from the CHARMM internal coordinate table, and (still keeping the unchanged atoms fixed) performing a vacuum energy minimization.

The average molecular surface of the 19 threaded sequences was computed using the "consensus volume" facility in the Grasp software.[14] Consensus volume produces a map (a 3D lattice of values) calculated by adding the value 1.0 to each grid point lying within the van der Waals volume of each molecule. For example, if there are five molecules and a grid point lies within all five, it will be assigned a value of 5.0. Some points will fall within the volume of only some molecules. This map can then be contoured at any level desired. For instance, contouring at the level of the total number of molecules will give the surface of the volume common to all molecules. For the 19 molecular structures, the contoured level of 5.0 is a good approximation for the surface average and was used here.

The electrostatic field was calculated for each of the 19 threaded structures with the internal linearized Poisson−Boltzmann solver of Grasp.[14] The electrostatic map was computed for every molecular structure and then averaged at each point of the map. The average electrostatic potential calculated at the molecular surface is shown in Results. Insight II software (Accelrys) was used to map residue hydrophobicity onto the solvent-accessible surface using the Engleman−Steitz hydrophobicity scale,[15,16] which is based on experimental and theoretical considerations of how well each residue enters a lipid bilayer from an aqueous environment.

The 11 most variant sequences of the original 19 were used in MD simulations using CHARMM, version 27.b2, with the all-atom force field, version 22.0.[17] The water model employed was a TIP3P (modified as in current use in CHARMM).[18] The electrostatic interactions were treated using the Ewald summation technique.[19,20] The peptide was placed in a solvent-filled box with the dimensions of 53 × 40 × 40 Å^3 and containing 2873 water molecules and 26 counterions (sodium and chloride). The 11 most variant sequences have different numbers of atoms and different total charges. To compensate for this, the number of water molecules and the number and composition of the counterions were slightly different for each sequence. The simulations were performed using periodic boundary conditions and using the Langevin piston constant pressure algorithm.[21] The simulation consisted of 300 steps of energy minimization using the steepest descent algorithm, followed by a further 5000 steps using the Newton−Raphson algorithm. The minimized systems were heated to 300 K, rescaling the velocities, for 30 ps. After that, an equilibration was performed for 1000 steps. During the dynamics simulations, the bond lengths involving hydrogen atoms were constrained, allowing the use of a 2 fs time step. The simulations were performed at constant temperature and pressure. The following parameters were employed during the NPT dynamics simulations: the mass of the pressure piston was 100 amu, the Langevin piston collision frequency was 10 ps^{-1}, and the thermal piston mass was 250 kcal ps^2. An isotropic pressure of 1 atm was applied, and the temperature was kept constant at 300 K. The simulations were run for 5 ns for each sequence.

Acknowledgment. This work was supported in part by a grant from Helmholtz-Gemeinschaft.

Supporting Information Available: Average backbone structure in pdb format. This material is available free of charge via the Internet at http://pubs.acs.org. The file containing the coordinates of the electrostatic surface readable by the GRASP software is available from the authors.

References

(1) Oldstone, M. B. A. *Viruses, Plagues and History*; Oxford University Press: New York, 1998.
(2) Zinkernagel, R. M.; LaMarre, A.; Ciurea, A.; Hunziker, L.; Ochsenbein, A. F.; McCoy, K. D.; Gehr, T. Neutralizing antiviral antibody responses. *Adv. Immunol.* **2001**, *79*, 1−53.
(3) Hahn, B. H.; Gonda, M. A.; Shaw, G. M.; Popovic, M.; Hoxie, J.; Gallo, R. C.; Wong-Staal, F. Genomic diversity of the acquired immune deficiency syndrome virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4813−4817.
(4) Wong-Staal, F.; Shaw, G. M.; Hahn, B. H.; Salahuddin, S. Z.; Popovic, M.; Markham, P. D.; Redfield, R.; Gallo, R. C. Genomic diversity of human T-lymphotropic virus type III (HTLV-III). *Science* **1985**, *229*, 759−762.
(5) Modrow, S.; Hahn, B. H.; Shaw, G. M.; Gallo, R. C.; Wong-Staal, F.; Wolf, H. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *J. Virol.* **1987**, *61*, 570−578.
(6) Check, E. Back to "plan A". *Nature* **2003**, *423*, 912−914.
(7) Burton, D. R.; Ryati, J.; Koduri, R.; Sharp, S. J.; Thornton, G. B.; Parren, P. W.; Sawyer, L. S.; Hendry, R. M.; Dunlop, N.; Nara, P. L. Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. *Science* **1994**, *266*, 1024−1027.
(8) Reed, J.; Kinzel, V. A conformational switch is associated with receptor affinity in peptides derived from the CD4-binding domain of gp120 from HIV I. *Biochemistry* **1991**, *30*, 4521−4528.
(9) Daura, X.; van Gunsteren, W. F.; Mark, A. E. Folding−unfolding thermodynamics of a beta-heptapeptide from equilibrium simulations. *Proteins* **1999**, *15*, 269−280.
(10) Hollaway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J. A priori prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J. Med. Chem.* **1995**, *38*, 305−317.
(11) Mihailescu, D.; Smith, J. C.; Reed, J. Solution structure of a putative HIV1 immunogenic peptide: computer simulation of the principal CD4 binding domain of gp120. *J. Med. Chem.* **2002**, *45*, 1019−1025.
(12) Mihailescu, D.; Reed, J.; Smith, J. C. Convergence in peptide folding simulation: multiple trajectories of a potential AIDS pharmacophore. *Biopolymers* **2003**, *70*, 121−133.
(13) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(14) Nicholls, A.; Sharp, K. A.; Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **1991**, *11*, 281−296.
(15) Engelman, D. M.; Steitz, T. A. The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. *Cell* **1981**, *23*, 411−422.
(16) Engelman, D. M.; Steitz, T. A.; Goldman, A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* **1986**, *15*, 321−353.
(17) MacKerell, J. A. D.; Bashford, D.; Bellot, M.; Dunbrack, R. L., Jr.; Evenseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhorn, B.; Reiher, I. W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanbe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(18) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. N.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(19) Essman, U.; Perea, I.; Berkowitz, M.; Darden, T.; Lee, H.; Pederson, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(20) Feller, S. E.; Pastor, R. W.; Rojnuckarin, A.; Bogusz, S.; Brooks, R. R. Effect of electrostatic force truncation on interfacial and transport of water. *J. Phys. Chem.* **1996**, *100*, 17011−17020.
(21) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613−4621.

JM031125S