Published in final edited form as:

J Mol Biol. 2009 January 16; 385(2): 675–691. doi:10.1016/j.jmb.2008.10.089.

***In silico* vaccine design based on molecular simulations of rhinovirus chimeras presenting HIV-1 gp41 epitopes**

Mauro Lapelosa^{1,2,3}, Emilio Gallicchio^{1,2}, Gail Ferstandig Arnold^{2,3}, Eddy Arnold^{2,3}, and Ronald M. Levy^{1,2}

¹ *BioMaPS Institute for Quantitative Biology, Rutgers University, Piscataway, New Jersey 08854*

² *Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854*

³ *Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, New Jersey 08854*

Summary

A cluster of promising epitopes for the development of HIV vaccines is located in the membrane-proximal external region (MPER) of the gp41 subunit of the HIV envelope spike structure. The crystal structure of the peptide corresponding to the so-called ELDKWA epitope (HIV-1 HxB2 gp41 residues 662–668) in complex with the corresponding broadly neutralizing human monoclonal antibody 2F5 provides a target for structure-based vaccine design strategies aimed at finding macromolecular carriers able to present this MPER-derived epitope with optimal antigenic activity. To this end, a series of replica exchange molecular dynamics computer simulations was conducted to characterize the distributions of conformations of ELDKWA-based epitopes inserted on a rhinovirus carrier and to identify those with the highest fraction of conformations able to bind 2F5. The length, the hydrophobic character, and the precise site of insertion were found to be critical for achieving structural similarity to the target crystal structure. A construct was obtained with a high degree of complementarity to the corresponding determinant region of 2F5. This construct was employed to build a high-resolution structural model of the complex between the 2F5 antibody and the chimeric HRV14:HIV-1 ELDKWA virus particle. Additional simulations, conducted to study the conformational propensities of the ELDKWA region in solution, confirm the hypothesis that the ELDKWA region of gp41 is highly flexible and capable of assuming helical conformations, as in the post-fusion helical bundle structure, as well as β -turn conformations, as in the complex with the 2F5 antibody. These results also suggest that the ELDKWA epitope can be involved in intramolecular and likely intermolecular hydrophobic interactions. This tendency offers an explanation for the observation that mutations decreasing the hydrophobic character of the MPER in many cases result in conformational changes that increase the affinity of this region for the 2F5 antibody.

Keywords

Vaccine design; Monoclonal antibody 2F5; Replica exchange method; Chimeric virus

Introduction

Due to the difficulty of providing universal anti-HIV drug therapies, the development of an effective AIDS vaccine remains the most promising long-term strategy to combat HIV/AIDS.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

There is a strong need for novel alternative approaches aimed at discovering an effective AIDS vaccine.¹ This work focuses on the development of vaccine components capable of eliciting broadly neutralizing immune responses. Optimized immunogens of this type could be combined with other B-cell immunogens as well as with T-cell immunogens for a multi-pronged attack on HIV, analogous to the multi-pronged attacks provided by drug cocktails. Binding of a broadly neutralizing antibody (nAb) to a pathogen containing large sequence diversity, such as HIV, is undoubtedly a prerequisite for the development of optimal B-cell immunogens. Nonetheless, the best immunogens to date remain insufficient for vaccination purposes.²

Antibodies can neutralize HIV-1 either by binding the virus surface and preventing its ability to interact with cellular receptors or by binding after virion attachment and blocking the subsequent steps involved with virus entry.³ The HIV-1 virus fuses to host cells with its surface spikes, composed of trimeric complexes of envelope glycoproteins gp41 (transmembrane) and gp120 (surface). gp120 subunits, decorated with host cell carbohydrates, form much of the external surface of the spike, which is anchored to the viral membrane via non-covalent association with the gp41 subunits. Binding of gp120 to the CD4 receptor of the target cell is the first step in the viral entry process, followed by a large conformational reorganization of the spike, eventually leading to viral fusion with the host cell through juxtaposition of the viral and host cell membranes.⁴ gp41 plays a major role in this process by forming a fusogenic six-helix bundle structure that is believed to drive the fusion process.⁵ gp41 contains ~ 345 residues⁶ and it is divided into an ectodomain, a membrane-spanning segment and a cytoplasmic tail.^{7; 8} The fusogenic six-helix bundle structure is formed when the N-terminal heptad repeats (NHR) and C-terminal heptad repeats (CHR) of the trimer of ectodomains of gp41 are able to associate in the absence of gp120.⁹

In this study, we model epitopes derived from the conserved membrane-proximal external region (MPER) of gp41. The MPER region is composed of a tryptophan-rich segment, presumably promoting some degree of positioning at the interface between water and the lipid bilayer of the viral membrane.¹⁰ Although the MPER is generally thought to be extensively cloaked in the context of the envelope spike, three of the most broad and potent nAbs for HIV-1, 2F5,¹¹ 4E10¹², and Z13e1^{12; 13} have been shown to bind here in close proximity. The epitope sites in the MPER are thus considered some of the most promising targets for AIDS vaccine development.^{14; 15} However, despite their promise, nAbs targeting the MPER, including 2F5, have been hypothesized to occur naturally only in a small minority of infected individuals. This phenomenon could be artifactual, being based on poor binding of serum antibodies to recombinant envelope mimics that may not have native-like conformations, and/or it could be that poor accessibility of the 2F5 epitope within the viral spike protein complex hinders recognition by the immune system or, alternatively, that this region could be exposed, but presented in a favorable immunogenic conformation only transiently.

The Arnold laboratory has developed combinatorial libraries to display HIV-1 epitopes connected to surface loops of human rhinovirus type 14 (HRV14)^{16; 17} with the aim of enhancing the probability of identifying epitope presentations able to evoke protective immune responses. HRV14, like all picornaviruses, has an icosahedral capsid shell consisting of 60 copies of the so-called protomeric subunit, comprising one copy each of the viral coat proteins VP1-VP4. Genetic engineering of HRV has resulted in the presentation of foreign immunogens able to trigger immunological responses specifically directed to the corresponding pathogenic antigens, in some cases, quite broadly.^{16; 17; 18}

Our efforts to develop chimeric HRV displaying the 2F5 epitope of HIV-1 gp41 (called ELDKWA after its core sequence) have provided indications that it may be possible to elicit protective immune responses with this approach. We hypothesize that stronger immunogens

in this class can be produced that are more likely to present the ELDKWA epitope in a conformation compatible with effective binding with the complementarity-determining region (CDR) of 2F5. The primary aim of the work described here is to apply computational modeling to identify chimeric HRV constructs which display preferentially the epitope in the same β -turn conformation as seen in the crystal structure in complex with 2F5. Although structure-based¹⁹ and bioinformatics²⁰ techniques have been used to find immunologically active sequences, the potential of computational modeling for vaccine design has only recently begun to be explored. Computational modeling approaches have been employed to design new peptides exhibiting molecular mimicry for known antigens;²¹ to characterize chimeric viruses presenting a double epitope with immunologically stimulating activity;²² and to compare free and bound crystal structures of monoclonal antibodies, illustrating the extent of induced fit effects on antibodies binding to their targets.²³

In this work, we seek to optimize the affinity of the 2F5 nAb to the ELDKWA epitope using a molecular simulation strategy similar in spirit to previous work on a bacterial epitope;²¹ the goal is to design a virus construct that maximizes the propensity of the epitope to assume the conformation seen in the known structure of the complex with the 2F5 Fab.²⁴ This approach assumes that the binding affinity of the ELDKWA epitope for 2F5 is increased when preferentially presenting this epitope in a conformation predisposed for binding. Conformational reorganization is an important factor for protein recognition and binding. Both binding partners incur a free energy penalty for reorganizing the ensembles of conformations present in their unbound forms to those compatible with complexation.²⁵ The binding affinity is enhanced when the binding partners already assume binding-competent conformations in their unbound forms. Minimizing the binding reorganization free energy is particularly important in vaccine design applications because, in general, the epitope composition cannot be freely varied to increase the binding affinity to target antibodies.

Achieving this goal requires high quality atomic resolution models. We employ in this study an energy function based on the OPLS-AA all-atom force field^{26; 27} combined with the AGBNP implicit solvent model,²⁸ which has been shown to be suitable for high resolution modeling of peptides and proteins.^{29; 30; 31} This free energy model has been recently reported to perform well for structure prediction of protein surface loops,³² a problem intimately related to the vaccine design strategy outlined here.

Another important requirement for this study is the use of a canonical conformational sampling method capable of predicting the thermodynamic stability of conformational states at physiological temperature. Here we employ the temperature replica exchange molecular dynamics (REMD) conformational sampling algorithm,³³ an advanced technique that consists of simulating many instances of the system in parallel over a series of temperatures. This canonical algorithm enhances conformational sampling efficiency alleviating some of the ergodicity problems associated with standard molecular dynamics algorithms, but it requires more extensive computational resources for its implementation.

We have carried out REMD simulations of various sequences derived from the MPER of gp41 in different structural contexts; as free peptides in solution as well as covalently attached within a solvent-accessible loop on the surface of HRV14. The peptide simulations provided information on the intrinsic conformational propensities of the epitopes, and suggested sequences that are more likely to be presented on HRV14 in the appropriate conformation for antibody binding. In addition, the peptide simulations provided insights into the native structure of the MPER within the context of the gp41 subunit. The simulations confirmed the hypothesis that the ELDKWA region of the MPER is highly flexible and capable of assuming helical conformations, as seen in the crystal structure of the post-fusion helical bundle that extends to the EL of the ELDKWA epitope,^{34; 35} and also β -turn conformations, as seen in the crystal

structure of an ELDKWA peptide complexed with the 2F5 Fab.²⁴ The results also provide an interesting perspective on the role of the hydrophobic interactions among the several tryptophan residues present in this region in influencing the structure of the MPER of gp41.

The simulations of the chimeric HRV carrying ELDKWA-derived peptides show that the foreign insert could fold in a variety of conformations, most of which are unsuitable for binding to the 2F5 nAb. We identified several energetic and topological factors that affect the population of binding-competent conformations of the inserted peptide and, using this knowledge, we identified a particular construct capable of displaying the ELDKWA epitope in a well exposed conformation able to fit into the 2F5 nAb cleft. The quality of agreement between some of the modeled conformations of the inserted peptide with the crystal structure of the 2F5 Fab complexed with the ELDKWA epitope has allowed us to produce a high-resolution structural model of the complex between the 2F5 nAb and the optimized chimeric HRV.

Results and Discussion

Limited exposure of the critical DKW residues of the ELDKWA epitope in the reference chimeric virus construct (cHRV1)

We have recently characterized an HRV14:HIV-1 chimeric rhinovirus displaying the gp41 ELDKWA epitope named cHRV1 (Table 1). When used to immunize guinea pigs, this chimeric virus was unable to elicit neutralizing anti-HIV-1 antibodies by itself; however, when the virus was boosted with a KLH-conjugated 14-mer peptide corresponding to the ELDKWA region (also unable to elicit neutralizing anti-HIV-1 antibodies by itself), antibodies capable of neutralizing diverse isolates of HIV-1 were elicited (data not shown). We speculate that the presentation of the ELDKWA epitope on the chimeric virus was not sufficient to elicit a broadly neutralizing response by itself, but that the peptide boosts stimulated a maturation of the secondary antibodies, enabling some of them to recognize and neutralize diverse isolates. Presumably, such a response reflects the fact that the chimeric virus is an imperfect immunogen, but not extremely so, as other virus immunogens with peptide boosts remained unable to elicit favorable immune responses.

To better understand the nature of this immunogen and to guide improved designs of immunogen constructs, we performed a replica exchange molecular dynamics (REMD) simulation of cHRV1 to assess the likely exposure and conformational propensities of the ELDKWA epitope on the surface of this rhinovirus. The simulation indicated that the ELDKWA epitope was highly mobile on the surface of the chimeric rhinovirus; however, the exposure of the inserted sequence was found to be rather limited on average (Fig. 1A) and, thus, not optimal for binding to the 2F5 nAb. The ensemble of conformations from the simulation was analyzed by hierarchical clustering based on pairwise C α RMSD's. In the most representative conformation, the side chains of the critical DKW residues point away from each other rather than being aligned as in the structure of the complex with 2F5,²⁴ and they fold over the VP2 subunit (Fig. 1A). In the dominant conformational cluster, only 17% of the PCG-ALDKWAS-SPDCS insert is solvent-exposed. This conformation has numerous clashes when superimposed with the corresponding paratope of 2F5. Nevertheless, in a minority of conformations, the critical DKW residues recognized by 2F5 are more solvent exposed; consistent with the observation¹⁹ that nAb 2F5 can bind and neutralize this virus. In Fig. 2A, the distribution of conformations in the simulated ensemble is shown; C α RMSD values of the ALDKWAS motif of each conformation are plotted with respect to that of the solution NMR structure of an ELDKWA-derived 13-mer peptide (1LCX)³⁶ and from the crystal structure of a MPER-derived 7-mer peptide in complex with 2F5 (1TJG).²⁴ The majority of conformations from the simulated ensemble of cHRV1 structures lie in the upper-right quadrant of Fig. 2A, indicating that they do not generally resemble the structural target represented by the crystal

structure of the corresponding peptide complexed with the 2F5 Fab²⁵ or the solution NMR structure³⁶.

The fraction of conformers of the inserted ELDKWA epitope dissimilar to the corresponding peptide in the crystal structure is a measure of the unfavorable reorganization free energy that must be overcome in order to achieve successful binding to the 2F5 nAb. As seen in Fig. 2A, the REMD trajectory at 310 K shows that this fraction is relatively large (approximately 95% as measured by clustering analysis), leading to the prediction that cHRV1 has a relatively low affinity for 2F5 due to the unfavorable reorganization free energy required for binding.

Higher affinity binding can be achieved by maximizing the fraction of binding-competent conformations. To achieve this goal we set out, as described next, to analyze the intrinsic conformational propensities of gp41-derived peptides as well as the influence of the framework provided by the HRV coat protein.

Dominance of α -helical and β -turn conformers of the ELDKWAS peptide sequence

To discriminate the effect of the intrinsic conformational propensity of the sequence from the influence of the framework in which it is embedded, we performed REMD simulations of the 7-residue ELDKWAS peptide in solution. Studies conducted by two laboratories^{36; 37} on related peptides did not provide a consensus interpretation of the solution NMR data. Unlike the same sequence integrated within gp41,³⁸ the free 7-residue peptide binds weakly to 2F5.³⁹ It has been further shown that longer 13–15 residue peptides containing the minimal ELDKWA epitope display increased binding affinity for 2F5.³⁷ These findings illustrate how the conformational propensity of this sequence can be modified by altering the surrounding structural context. They suggest that, in longer sequences, the fraction of the epitope core sequence (DKW) compatible with 2F5 binding can be increased relative to that of the ELDKWAS minimal peptide, and also that some of the surrounding residues (such as the adjacent leucine residues in LELDKWASL) may make favorable contacts to stabilize binding with 2F5.

REMD simulations of the minimal 7-residue peptide ELDKWAS indicate that this peptide is characterized by significant structural heterogeneity. Fig. 3 presents the computed distribution at 313 K of the C α RMSD values of the ELDKWAS motif with respect to the 1LCX NMR structure³⁶ and the 1TJG crystal structure²⁴ as described above for the cHRV1 chimera. The 1LCX solution NMR structure is the 3_{10} helical structure reported by Biron et al.³⁶ of a 13-residue peptide containing the ELDKWAS motif (Barbato et al.³⁷ reported a less homogeneous NMR ensemble for a closely related peptide), while the 1TJG crystallographic structure represents the β -turn conformation of the ELDKWAS peptide complexed with the 2F5 Fab reported by Ofek et al.²⁴

Hierarchical clustering analysis based on pairwise C α RMSD's of the simulated ensemble identifies two highly populated clusters corresponding to helical (Cluster 1) and type I β -turn (Cluster 2) conformations, representative of the solution NMR structure and the crystal structure, respectively. The “helical” cluster (Cluster 1), which corresponds approximately to the samples in the lower right quadrant of figure 3, has a fractional population of 47%. The structural correspondence between the representative structure of Cluster 1 and the NMR solution structure is illustrated in Fig. 4A. In contrast, the structure representative of Cluster 2 is more similar to that of the crystal structure (1.5 Å RMSD) than the NMR structure (2.4 Å RMSD). This “ β -turn” cluster (Cluster 2), which corresponds approximately to the samples in the upper left quadrant of figure 3, has a fractional population of 44%. The structural similarity between the representative structure of Cluster 2 and the crystal structure from the 2F5 complex is illustrated in Fig. 4B. Here both the crystal and modeled structures adopt a type I β -turn conformation. The difference between the conformations in Cluster 2 and the crystal structure

is mainly limited to the orientation of the W666 indole ring, which is often rotated toward the interior of the peptide in the simulated conformations rather than facing outward as in the crystal structure (Fig. 4B). This bias observed in the computed structural ensemble was confirmed by performing a molecular dynamics (MD) simulation starting from the conformation of the peptide in the crystal structure in complex with 2F5. In this simulation, the side chain of the W666 rotated after 8 ns from the outward orientation to the inward orientation as observed in most of the conformations from Cluster 2.

These results indicate that the structure of the ELDKWAS peptide interconverts in solution between an α -helical conformation and a type I β -turn conformation. This structural heterogeneity is presumably the origin of the discordant conclusions of NMR studies in which this peptide has been described as assuming a 3_{10} helix conformation³⁶ or multiple diverse conformations.³⁷ Both α -helical and β -turn conformations have been observed, corresponding to different biological functions for this region: the immunogen has been seen to interact with 2F5 in a β -turn conformation²⁴ and to interact (at least in its N-terminal portion) with the corresponding oligomeric MPER gp41 segments in the fusion-driving six-helix bundle structure.^{34; 35} Although it is not known if the six-helix bundle or the triple helix that is thought to precede it actually extends all the way into the ELDKWA epitope, it is considered likely to do so, especially given that: (1) ELDKWA-based peptides can form α -helical conformations in aqueous solution;³⁶ (2) the epitope for 4E10 (distal to the ELDKWA epitope with respect to the six-helix bundle) is known to form an α -helix in the presence of the antibody; and (3) there are no major helix-breaking residues within the ELDKWA epitope. In addition, (4) related helical conformations were obtained in an NMR study of a 19-residue peptide incorporating the ELDKWAS sequence in dodecyl phosphocholine micelles, mimicking the viral membrane environment.⁴⁰ In this study, the peptide was shown to be positioned at the interfacial region between the membrane and the solvent, with the tryptophan rings preferentially inserted into the membrane bilayer.

It has been shown that alanine substitutions in the gp41 MPER of a HIV pseudovirus (HIV-1_{JR2}) can alter antigenic recognition by the 2F5 nAb as well as by the 4E10 nAb, whose epitope is proximal to the 2F5 epitope.³⁸ In particular, it was observed that alanine single mutations of the DKW core of the 2F5 epitope (D664A, K665A, and W666A) substantially lowered the affinity of 2F5 for the pseudovirus. In contrast, the replacement by alanine of many hydrophobic residues nearby, including those in the 4E10 epitope (W670A, W672A, and W678A), increased the binding affinity for the 2F5 nAb, even though these residues have not been shown to have direct binding interactions with the 2F5 nAb. Similarly, alanine substitution for W666 in the center of the ELDKWA epitope was found to increase the binding affinity of 4E10 to its epitope, also in the absence of any evidence for direct binding interactions. Thus, the affinity of the 2F5 and 4E10 nAbs for the MPER of gp41 is influenced by the presence of hydrophobic residues even though these may not be directly involved in binding and are relatively distant in sequence from the characterized epitopes. These observations suggest that the composition of the hydrophobic flanking regions indirectly influences the binding affinity by modulating the conformation of the epitopes at a distance. One possibility is that reducing the hydrophobicity of these flanking regions could favor conformations of the epitope regions with better shape complementarity with the corresponding paratope on the antibody molecules.

These observations led us to explore the possibility that the affinity of the ELDKWAS sequence to 2F5 could be determined by the conformational propensity of the ELDKWAS sequence to assume a β -turn conformation. To test this possibility, we performed *in silico* alanine scanning REMD simulations aimed at predicting the conformational propensities of a series of peptides derived from the native ELDKWA epitope (ELDKWAS, Peptide 1) by mutating each residue in turn to alanine (Peptides 2–7). We found that the average helicity of the overall ensemble (35.1%, 37.2%, and 38%, respectively) of the peptides with the D664A, K665A, and W666A

mutations were significantly greater than that of both natural sequences (20% for ELDKWAS and 15% for ALDKWAS). These are the same mutations that were shown to most markedly decrease the affinity of 2F5 to the envelope spikes of the corresponding pseudoviruses.³⁸ The observation that constructs in which the ELDKWAS sequence was incorporated into helical frameworks have no detectable 2F5 reactivity,⁴¹ and the present computational predictions indicate that, in addition to the loss of critical binding interactions, the stabilization of helical conformations of the 7-mer epitope sequences by the alanine replacements of the DKW further disfavors 2F5 binding.

Reduction of hydrophobic residue content enhances epitope exposure in a 21-mer MPER-derived peptide

To further investigate the role of alanine substitution on the binding affinity of 2F5 as well as on the binding affinity of 4E10, the most broadly nAb known against HIV, which binds between ELDKWA and the viral membrane,³⁸ we conducted solution REMD simulations of an MPER-derived peptide of 21 amino acids in length, containing the epitopes of both 2F5 and 4E10 (sequence LLELDKWANLWNWFDISNWLW from the HIV-1 HxB2 gp41 residues 660–680, with the epitopes underlined). The results of the RMSD analysis of the 313 K simulated ensemble based on the ELDKWAN core motif with respect to the crystal and NMR reference structures used above (Fig. 5A) shows similar grouping to that seen for the ELDKWAS peptide (Fig. 3). The helical and β -turn conformations correspond to 14% and 25% of the structures, respectively. In addition, we find a significant number (61%) of conformations (upper right in Fig. 5A) in which the ELDKWAN motif does not assume either helical or β -turn structures. Some of these conformers are likely to correspond to a series of dominant conformers identified by clustering analysis that all display extensive long-range hydrophobic interactions, as in the representative structure shown in Fig. 6.

The large hydrophobic nucleus observed in these structures is mainly stabilized by Trp-Trp π -stacking interactions. The critical W666 and W672 residues of the 2F5^{12; 24} and 4E10^{12; 42} epitopes, respectively, appear to play central roles in the formation of the hydrophobic nucleus in nearly all of the physiological temperature structures. Thus, if such intramolecular hydrophobic interactions occur *in vivo* or if alternative intermolecular hydrophobic interactions occur instead *in vivo* among the conserved tryptophans, then these tryptophans would be unavailable for intermolecular binding interactions with their cognate antibodies (or possibly other biological partners). These results indicate that the binding affinity of the gp41 MPER is modulated by a competition between conformations with extensive intramolecular hydrophobic interactions in which W666 and W672 are likely to be sequestered and conformations in which these residues are available for binding. The greater antigenic affinity observed for alanine-substituted mutants is rationalized by the tendency of these mutations to disrupt the hydrophobic interactions, thereby shifting the equilibrium toward conformers in which the W666 and W672 side chains become more available for binding to their respective antibodies. We confirmed this hypothesis by carrying out a REMD simulation of the quadruple mutant peptide, LLELDKWANLANAFDISNALA, in which four tryptophan residues (W670, W672, W678, and W680) of the natural MPER 21-residue peptide were simultaneously replaced with alanine residues. We found that in the conformational ensemble of this mutated sequence, the tendency toward hydrophobic nucleation was considerably diminished relative to the tryptophan-rich natural sequence. Concomitantly, there was a greater occurrence in this ensemble relative to the ensemble of the natural sequence of conformations in which the backbone of the DKW sequence matched structurally with that of the same residues in the crystal structure in which a peptide is complexed with the 2F5 nAb [35% for the quadruple alanine mutant (Fig. 5B) vs. 19% for the natural sequence (Fig. 5A)]. The clustering analysis showed that conformations structurally similar to the crystal structure (1TJG) became more populated and the population of the cluster of helical conformations decreased (Figs. 4A and

4B). The REMD results are consistent with a small free energy difference between two functionally relevant conformations. Furthermore, in all of the conformations of the alanine-substituted peptide, the critical W666 residue did not form hydrophobic contacts and was fully available for intermolecular interactions (*e.g.*, with 2F5). This is in contrast to the natural peptide ensemble, where even for those of conformations (19%), in which the DKW motif matched the structure of the peptide complexed with 2F5, W666 was sequestered and not available for binding. This is illustrated in Table 2, which reports a quantitative measure of the degree of pairwise hydrophobic packing for the five key tryptophan positions (W666, W670, W672, W678, and W680) in the natural 21-mer sequence among conformations in which the structure of the DKW motif matched that of the corresponding residues in the complex with 2F5. In this table, values near one indicate a high degree of packing whereas values near zero correspond to little or no packing. The data show that, in the natural peptide, the critical W666 residue forms extensive hydrophobic contacts, particularly with W672 and W670 (with packing degrees of 0.77 and 0.72, respectively) and also interacts with W678 and W680 despite their relatively distant sequence position. Similar trends toward the formation of extensive hydrophobic contacts are observed for the W672 residue, which is part of the 4E10 epitope. In contrast, the same analysis conducted on the conformational ensemble corresponding to the alanine-mutated sequence resulted in values of the degree of pairwise packing close to zero (not shown), indicating that, in this peptide, the occurrence of intramolecular hydrophobic interactions is greatly reduced.

Although the simulation performed here suggests that intramolecular hydrophobic aggregation hinders exposure of the 2F5 and 4E10 epitopes, hydrophobic associations within the gp41/gp120 trimeric complex and/or with the viral membrane or other biological partners could similarly make them less available for interacting with their cognate antibodies. In particular, the tryptophan residues, whose substitution by alanine promoted improved binding by the antibody to the adjacent MPER epitope, may normally insert themselves in the viral membrane. In possible support of this hypothesis, the CDR-H3 fingers of 2F5 and 4E10 could interact with more than the MPER epitopes thus far characterized, including possibly the viral membrane.^{24; 42} Nevertheless, the increase in the exposure of the 2F5 and 4E10 epitopes observed in the simulations upon reduction of the hydrophobic character of the 21-mer MPER-derived peptide is likely to occur regardless of which specific hydrophobic sequestering mechanism might be operating.

The simulation indicates that the natural 21-mer MPER-derived peptide is highly flexible, consistent with the hypothesized large conformational reorganization that occurs in the MPER during the viral fusion process. One of the most populated conformers adopts an α -helical conformation, consistent with the conformation of the crystal structure of the late-fusion state of gp41 and hypothesized to play a major role in the viral entry process.^{37; 43; 44} During the fusion process, it is believed that the outer layer of the six-helix bundle, corresponding to the MPER of gp41, makes extensive contacts with the viral membrane, mediated by the numerous tryptophan residues present in this region. We indeed observed that in the helical conformers (Cluster 1) obtained from the simulation of the 21-mer natural peptide, the side chains of the aromatic residues (especially W666, W670, and W672) are often located on the same side of the helix, in an orientation consistent with the hypothesized propensity of these residues to be localized at the membrane-protein interface.⁴⁵

Insertion of the quadruple alanine-substituted 21-mer MPER-derived peptide into HRV (cHRV2) results in poor exposure of the epitope

Longer MPER sequences are more likely to reproduce the natural conformation of the MPER as part of the HIV spike complex. Simulations of the structures of the quadruple alanine-substituted 21-mer peptide (LLELDKWANLANAFDISNALA) in solution resulted in an

increase in the exposure of the critical W666 side chain and an increase in the fraction of β -turn conformations of the DKW region of the epitope (characteristic of the crystal structure of the complex with 2F5). In an attempt to exploit both of these potential benefits in the design of the chimeric virus, we decided to model a chimeric HRV (cHRV2) in which the 21-mer alanine-mutated peptide was inserted into the VP2 puff of the parent HRV14 with removal of the VP2 residues S158-A159-N160-E161, as was done with cHRV1. The rationale behind the alanine substitutions was to avoid the formation of the intramolecular hydrophobic clusters involving W666. Further analysis of the REMD trajectory, however, showed that the majority of the conformations of the engineered loop were not compatible with 2F5 recognition (Fig. 2B). The other shortcoming we observed is the relatively limited exposure of the region of the engineered loop critical for 2F5 binding (Fig. 7). This is likely due to the tendency of the loop to fold over onto the HRV capsid, presumably due to the length of this hydrophobic insert. This result indicates that although the quadruple mutant peptide often adopts a favorable conformation for antibody binding in solution, when linked to HRV it adopts a less favorable conformation.

The structure of the 7-mer ELDKWAS peptide inserted onto HRV (cHRV3) matches that of the peptide in the 2F5 antibody complex

In an attempt to restructure the DKW core epitope and to increase the likelihood of best exposing this critical region at the insert/capsid interface, we reduced the length of the insert. To this end, we modeled a series of shorter engineered loops, the most promising of which resulted in a third chimeric construct (cHRV3). In this construct, 7 residues of the 2F5 epitope (ELDKWAS) were inserted into the VP2 puff of HRV14. In contrast to cHRV1, this chimeric construct was allowed to retain the four HRV14 VP2 residues that were previously removed (S158, A159, N160, and E161) and the insert was placed between N160 and E161.

The REMD simulation of cHRV3 showed that the engineered loop often adopts a structure likely to be recognized by 2F5 (Fig. 2C). As many as 41% of the modeled conformations the ELDKWAS core epitope are within 2 Å C α RMSD of the crystal structure of the corresponding peptide complexed with the 2F5 Fab. In these conformations the ELDKWAS insert protrudes from the body of the virus capsid, extending approximately 19 Å from base to tip (Fig. 1B). The loop contains 37 residues from VP2 (residues L156 to H193) and includes the 7-residue 2F5 core epitope sequence (ELDKWAS) at its center corresponding to the tip region of the loop. The engineered loop can be sub-divided into three structural regions: the N-terminal base, a flexible body extended away from the base, and the C-terminal edge which is folded within the base of the VP2 subunit. The REMD model suggests that a network of hydrogen-bonding interactions at the two bases of the loop may account for the favorable positioning of the ELDKWAS motif at the tip of the loop (Fig. 8). The carbonyl oxygen of E662 (for the inserted peptide, the residue numbers follow HIV-1 gp41 numbering, and the residues of VP2 conserve the wild-type numbering) hydrogen bonds with the hydroxyl group of S668 of the same loop (Fig. 8), and VP2 H135 forms hydrogen bonds at the N-terminal and C-terminal ends of the inserted peptide, with the carbonyl oxygen of N160 and the amide nitrogen of G164 respectively. Also, K167 forms hydrogen bonds with the carbonyl oxygen of A159 and with the O ϵ 1 of E161 at the two ends of the epitope.

Similar stabilizing interactions are not seen in the cHRV1 and cHRV2 constructs, which however are characterized by alternative interactions that prevent correct positioning of the core epitope. For example, in cHRV1 (Fig. 1A) the backbone of linker residue P158 of the N-terminal PCG linker region (numbered according to the VP2 wild-type numbering) upstream from the 7 residue ELDKWAS insertion is often hydrogen bonded to W666, causing it to fold inward. This proline residue also appears to induce an unfavorable backbone conformation of the engineered loop. Similarly, the amino hydrogen and carbonyl oxygen of C159 of the same

linker form hydrogen bonds with O ϵ 1 of S668 and H δ 1 of H135, respectively, limiting the exposure of the inserted epitope. This pattern of interactions causes a lengthening of the distance between the N-terminal and C-terminal ends of the engineered loop in cHRV1 relative to cHRV3 (approximately 9 Å in cHRV1 vs. 6 Å in cHRV3), thereby favoring more extended and less exposed conformations of the epitope in cHRV1.

In the simulation with cHRV3, we observe little global deviation of the system from the crystal structure of wild-type HRV14. Superimposition based on Ca atoms reveals a relative RMSD of 0.86 Å of the VP2 subunit between the representative simulated conformation and the crystal structure. Most of the displacement occurs at the insertion site, between VP2 residues N160 and E161 of the VP2 puff. It appears that the HRV carrier is stable, owing to the plasticity of the insertion region which accommodates the insertion without major disruptions of native interactions.

The DKW region of the epitope at the tip of the loop was found primarily to adopt a β -turn conformation (Fig. 1B). Previous work on a cyclic peptide incorporating a lactam bridge between the two proximal amino acids flanking the DKW segment⁴⁶ suggests that, even though the type-1 β -turn structural motif is the basic feature for recognition by 2F5, high binding affinity requires a specific cation- π stacking interaction between K665 and W666 which is also observed in our simulation. We have also analyzed the distribution of solvent-accessible surface areas (SASA) of the core DKW segment from the ensemble of conformations generated by the simulation. Comparison of the SASA distribution of cHRV3 with those of cHRV1 and cHRV2 (Fig. 7) shows that, in cHRV3, the foreign epitope has significantly greater exposure. This suggests that the epitope would have a higher probability of protruding deeply into the binding cleft of the concave paratope of 2F5. This is in contrast to the W666 side chain of the 7-mer ELDKWAS peptide in solution, which assumed a conformation that was rotated inward, with the side chain sequestered in an intramolecular hydrophobic core without significant reorganization of the predominantly β -turn backbone structure or the α -helical backbone structure (Figs. 4A and 4B). The occurrence of these differing structures, combined with the observed inward rotation of W666 in the 21-mer peptide (Fig. 5), suggests that there may be a common hinge mechanism that occurs during antibody recognition as well as during some of the other intermolecular interactions that characterize this dynamic region of gp41.

Structural model of the antibody/cHRV3 complex

Noting that the structure of the DKW segment of cHRV3 matched that of the corresponding segment in the crystal structure of the 2F5 complex in a significant fraction of conformations, we have constructed a structural model of the complex of cHRV3 with the 2F5 nAb by superimposing the ELDKWAS residues of a representative conformation of cHRV3 to the corresponding residues of the 1TJG crystal structure of the 2F5 complex (Fig. 9). The structure so obtained exhibits a remarkable shape complementarity between the modeled epitope and the 2F5 paratope from the crystal structure. The side chains of the DKW segment project inside the 2F5 hydrophobic crevice as they do in the crystal structure. The structure also indicates a high degree of charge complementarity between the paratope and the epitope and the buried surface is mainly composed of hydrophobic amino acids. The epitope occupies a cavity approximately 10 Å deep by 14 Å wide at the interface between the light and heavy chains of the antibody hypervariable region. The intermolecular contacts between cHRV3 and 2F5 include salt bridges and hydrophobic interactions. The lysine side chain of K665 points to a hydrophobic pocket and is proximal (4.3 Å) to D54 of the heavy chain. The D664 side chain forms interactions (3.9 Å) with R95 of the heavy chain. Packing interactions can also be observed between the rings of W666 and F32 of the heavy chain. The side chain positions and

relative orientations of K665 and W666, which are both critical for binding, are reproduced remarkably well in the modeled complex. The same is true for the D664 side chain.

Conclusions

We have used the replica exchange molecular dynamics conformational sampling method to study the conformational distributions of peptides and chimeric constructs derived from the immunogenic MPER of the gp41 protein of the HIV-1 envelope spike. The main aim of the work was to find chimeric constructs based on a rhinovirus carrier (HRV14) able to present the HIV-1 ELDKWA epitope with a high fraction of antibody binding-competent conformations. Such constructs are predicted to bind with high affinity to the 2F5 antibody and thus to have greater likelihood to elicit, in turn, the production of broadly neutralizing antibodies with similar characteristics. We have identified a promising HRV14:HIV-1 chimeric construct (cHRV3) obtained by inserting the ELDKWA sequence into the VP2 puff of HRV14. The predicted structure of the ELDKWA epitope of this construct matched that of the crystal structure of the corresponding peptide complexed with 2F5 with high frequency. The close similarity of many conformations generated by REMD to the crystal structure allowed us to create a structural model of the complex of the chimeric HRV14 with the 2F5 antibody. The binding-competent conformations of the epitope when inserted into HRV14 are stabilized by a network of specific hydrogen bonds as well as a number of specific hydrophobic interactions and by the absence of competing electrostatic and hydrophobic interactions that unfavorably affect the other constructs we have investigated.

The results of this work complement published work and ongoing biochemical and spectroscopic research aimed at structure-based vaccine design through the investigation of structural propensities of MPER-derived peptides in different environments^{37; 40; 47} including insertion onto protein carriers.^{18; 48; 49} The relevance of the present modeling study extends also to research aimed at uncovering the structural properties of the gp41 subunit as part of the HIV-1 envelope spike and the mechanisms by which the relevant structures elicit broadly neutralizing immunological activity. Our results confirm the hypothesis that the MPER of gp41 is highly flexible and capable of assuming a variety of conformations consistent with its biological function, and suggest that epitope exposure in gp41 is countered by intramolecular hydrophobic interactions. The latter observation offers an explanation for the finding³⁸ that mutations of hydrophobic residues distant in sequence and not directly involved in antibody binding, by weakening these hydrophobic interactions, can substantially increase the affinity of the 2F5 and 4E10 nAbs for gp41. The conformational heterogeneity of this region of gp41 and its interactions with components of the envelope spike and with the viral membrane likely play important roles in the processes of antibody recognition and antibody maturation.

The computational methodology introduced in this paper is generally applicable to the *de novo* design of recombinant vaccines when, such as in the present case, structural information of a target is available. By identifying the factors that influence the presentation of the epitope toward the target, modeling work can guide the selection of vaccine candidates. The physical insights concerning the conformational propensities of epitopes inserted onto HRV14 provided by this work will facilitate the designs of focused combinatorial libraries of HRV14-based chimeras. The main design principles that emerged from this study are to limit the size of the insert, to optimize the central positioning of the critical DKW motif within the insert, as well as to minimize the hydrophobic character and the structural constraints determined by the flanking sequence. These focused libraries will be assayed in terms of their affinity for the 2F5 nAb, and their ability to elicit strong and broadly neutralizing immune responses. This work is currently under way in our laboratory.

Materials and Methods

Replica exchange molecular dynamics

The temperature replica exchange molecular dynamics (REMD) method^{29; 33; 50} as implemented in the IMPACT computational package⁵¹ was used in this work to explore the conformational space of several peptides derived from the MPER of HIV gp41 in solution as well as inserted onto the HRV14 viral capsid. REMD is an advanced canonical conformational sampling algorithm designed to help overcome the sampling problem encountered in biomolecular simulations. It consists of running a series of MD simulations of the molecular system in parallel over multiple processors at different temperatures. Periodically a Monte Carlo move is attempted, aimed at swapping atomic coordinates of one replica with those of another. The acceptance probability expression for the swap move is derived based on the detailed balance condition. Typically the lowest temperature replica corresponds to the temperature of interest whereas the high temperature replicas provide a means to rapidly explore the extent of conformational space. In the temperature replica exchange method, efficient interconversion between low energy conformations occurs whereby the temperature switches from low to high, allowing the conformation to overcome potential energy barriers and then transition to another conformation. A new low energy conformation can then be established when this conformation transfers back to low temperature replicas.

Force-field and implicit solvation model

All simulations employed the OPLS-AA fixed charge all-atom force field^{26; 27} and the AGBNP²⁸ implicit solvent model to mimic the water environment. Simulations using implicit solvent models are not as computationally intensive as those using explicit solvent models; this is particularly true for REMD simulations because many fewer replicas are needed to achieve reasonable replica exchange rates due to the smaller system size and broader overlap of the energy distributions among replicas at different temperatures. AGBNP is an implicit solvent model based on a novel implementation of the pairwise descreening scheme of the generalized Born model^{52; 53} for the electrostatic component, and a novel non-polar hydration-free energy estimator. The non-polar term consists of an estimator for the solute solvent van der Waals dispersion energy designed to mimic explicit solvent solute-solvent van der Waals interaction energies in addition to a surface area term corresponding to the work required for cavity formation. It has been used to study protein conformational equilibrium and protein allostery.³¹

Peptide simulations

The REMD simulations in solution of gp41-derived peptides were started from the fully extended conformation. Before performing REMD simulations the system was equilibrated at each temperature. The temperature ladder was adjusted to obtain sufficient overlap of the potential energy distributions at neighboring temperatures to ensure frequent acceptance of temperature exchange moves. The REMD simulations of the ELDKWAS peptide and its alanine-substituted mutants were carried out using eight replicas for 20 ns (total production run 160ns) at the temperatures 313, 328, 345, 363, 381, 400, 421, and 442 K. Convergence progress was analyzed by monitoring the total potential energy and key interatomic distances as a function of time. Helical conformations were determined using the program STRIDE⁵⁴ for each residue in the trajectory of the 313 K replica. The average helicity is obtained by averaging the fractional helical populations of the residues in the sequence. REMD simulations of the LLELDKWANLWNWFDISNWLW peptide, and its mutants were similarly performed using replicas at 313, 328, 347, 367, 388, 411, 434, 460, 486, and 515 K.

Model preparation

The three HRV14:HIV chimeric constructs investigated in this work were obtained by insertion of gp41-derived epitopes in the β E- β F region of the so-called VP2 puff of the VP2 subunit of the HRV14 protomeric unit, the largest of three loops forming the neutralizing immunogenic site II (NIm-II) of HRV14.^{55; 56} The first chimera (designated as cHRV1 in this paper and as 14-C40-1 in Arnold et al., submitted) was obtained by removing four residues of the VP2 puff (S158, A159, D160, and N161) and replacing them with the “ELDKWA” epitope sequence (ALDKWAS) flanked by PCG on the amino-terminal side of the epitope and SPDCS on the carboxyl-terminal side of the epitope (Table 1). The second chimeric construct (cHRV2) was obtained by insertion of the sequence LLELDKWANLANAFDISNALA, a quadruple alanine-substituted MPER peptide corresponding to residues 660–680, in the same HRV site as for cHRV1 except for the absence of the flanking linkers, PCG and SPDCS (Table 1). The third chimeric construct (cHRV3) contains the sequence ELDKWAS inserted between the N160 and E161 of VP2, without the removal of any HRV14 residues or the addition of any flanking residues. Starting structures for the cHRV1, cHRV2, and cHRV3 chimeras have been obtained by homology modeling using the program Prime (Schrodinger Inc.) and the structure of wild-type HRV14 (PDB ID: 4RHV) as a structural template (97% sequence identity). These models can be built with a high degree of confidence since the target and template structures are expected to differ only in the proximity of the inserted epitope, an HRV region that is only 12 residues long. Therefore, the estimated confidence factor for the homology models was high. The energy-minimized models have been assessed for quality using ProCheck, indicating 84.6%, 14.6%, 0.7% and 0.1% in the most favored regions, allowed regions, generously allowed regions, and disallowed regions of the Ramachandran plot,⁵⁷ respectively.

Nine peptides were used for these studies (Table 1). All have NH_3^+ groups at their amino termini and COO^- groups at their carboxy termini. There were seven natural and singly alanine-substituted (corresponding to HxB2 gp41 residues 662–668) 7-mer peptides, and two natural and quadruply alanine-substituted (corresponding to HxB2 residues 660–680) 21-mer peptides,.

REMD simulations of the chimeric constructs

REMD simulations of the cHRV1, cHRV2, and the cHRV3 viral constructs were conducted to estimate the conformational propensities of the inserts. The production REMD runs consisting of 30 replicas with temperatures in the range 310K to 495K were 2ns per replica in length for a total of 60ns for each of the three systems. The production runs were preceded by minimization and equilibration calculations starting from the comparative models described above. To reduce computational complexity, the residues within all atoms further than 20 Å from any atom of the foreign insert were not included in the model. Because of the different sizes of the inserts, the number of residues completely free to move ranged from 227, 276, and 176 in cHRV1, cHRV2, cHR3 respectively. Portions of the VP1 and VP3 protomeric units within 20 Å of any atom of the insert of the viral capsid were included as well. The positions of Ca atoms of the residues at the ends of dangling protein chains were harmonically restrained using an isotropic force constant of 0.3 kcal/mol/Å². This setup prevents unfolding of the simulated protein region at high temperature while providing a reasonably accurate description of the protein environment surrounding the inserted sequence. This methodology was tested by applying it to the chimeric HRV displaying part of the HIV-1 V3 loop studied by Smith et al.¹⁷ We confirmed that the main cluster of conformations of the loop generated by REMD structurally matched that of the crystal structure.

Hydrophobic packing measure

An algorithm has been designed aimed at the quantitative characterization of the degree of packing interactions among residues. The measure is based on the ratio

$$x_{ij} = \frac{R_g(i, j)}{\sqrt{R_g(i)R_g(j)}}$$

between the radius of gyration, $R_g(i, j)$, of a residue pair and the geometric mean of the radii of gyration, $R_g(i)$ and $R_g(j)$, of the individual residues. A small value of x_{ij} approaching 1 indicates that the two residues are in contact. Conversely, a value of x_{ij} much larger than 1 indicates that the two residues are distant from each other relative to their sizes. A normalized degree of packing, $f(x_{ij})$, ranging between 0 (residues are not in contact) and 1 (residues are in contact) is obtained by applying the following linear interpolation function

$$f(x) = \begin{cases} 1 & x \leq x_{\min} \\ ax+b & x_{\min} < x < x_{\max} \\ 0 & \text{otherwise} \end{cases}$$

where

$$a = \frac{-1}{x_{\max} - x_{\min}}$$

and

$$b = \frac{x_{\max}}{x_{\max} - x_{\min}}.$$

The function above returns 1 for x_{ij} smaller than $x_{\min}$, 0 for x_{ij} larger than $x_{\max}$, and intermediate values for x_{ij} between these two limits. Here we set $x_{\min} = 2.5$ and $x_{\max} = 3.5$.

The degree of packing, $\langle f(x_{ij}) \rangle$, between residues i and j averaged over an REMD trajectory measures the fractional propensity of the two residues to be in contact. The closer this quantity is to 1, the higher the occurrence is of packing between a pair of residues. We used this measure to analyze the occurrence of intramolecular hydrophobic contacts in the simulation of peptides derived from the MPER of gp41. In that context we applied the packing measure only to those conformations in which the most essential ELDKWA residues, DKW, assumed a conformation similar to that obtained from the crystal structure, 1TJG, based on a 30° threshold maximum difference between the backbone dihedral angles.

Solvent-accessible surface calculation

The solvent-accessible surface area (SASA)⁵⁸ was determined to quantify the conformational exposure of the engineered loop insert on the chimeric HRV. The SASA has been calculated for each atom using the program SURFACE based on the algorithm of Sridharan et al.⁵⁹ with a water probe radius of 1.4 Å. Hydrogen atoms have been ignored in the calculation of the SASA. The distribution of SASA for the ELDKWA epitope residues DKW has been determined for the structures present in the replica exchange trajectories of the chimeric constructs cHRV1, cHRV2, and cHRV3.

Hierarchical clustering

The computed conformational ensembles have been organized into conformational clusters using the XCluster program (Schrodinger, Inc.). XCluster implements a hierarchical clustering algorithm⁶⁰ based on the minimal root mean square deviation (RMSD) between user-defined sets of atoms of two conformations. The clustering procedure starts by computing the all-to-all distance matrix between the selected sets of conformations. A series of clustering arrangements, or clustering levels, is then produced. At the first level each of the N conformations forms a separate cluster. The following level with $N-1$ clusters is obtained by joining into a single cluster the pair of conformations closest to each other. This process continues by joining at each step the closest pair of clusters until, at the N -th level, only one cluster encompassing the entire ensemble remains. XCluster employs the single-linkage convention to compute the distance between two clusters, defined as the smallest distance between any two conformations of the two clusters. Each clustering level i corresponds to $N-i$ clusters. The choice of the clustering level to pick is aided by clustering quality measures computed by XCluster such as the minimum separation ratio, which is the smallest ratio between the shortest intra- and inter-cluster distances among all clusters. Most of the results obtained here have been obtained at the clustering level corresponding to approximately ten clusters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work has been supported in part by NIH grant GM30580 (to RMC) and by NIH grant 1AI071874 (to GFA).

References

1. Klausner RD, Fauci AS, Corey L, Nabel GJ, Gayle H, Berkley S, Haynes BF, Baltimore D, Collins C, Douglas RG, Esparza J, Francis DP, Ganguly NK, Gerberding JL, Johnston MI, Kazatchkine MD, McMichael AJ, Makgoba MW, Pantaleo G, Piot P, Shao Y, Tramont E, Varmus H, Wasserheit JN. Medicine. The need for a global HIV vaccine enterprise. *Science* 2003;300:2036–9. [PubMed: 12829768]
2. Montero M, van Houten NE, Wang X, Scott JK. The membrane-proximal external region of the human immunodeficiency virus type 1 envelope: dominant site of antibody neutralization and target for vaccine design. *Microbiol Mol Biol Rev* 2008;72:54–84. [PubMed: 18322034]
3. Burton DR, Desrosiers RC, Doms RW, Koff WC, Kwong PD, Moore JP, Nabel GJ, Sodroski J, Wilson IA, Wyatt RT. HIV vaccine design and the neutralizing antibody problem. *Nat Immunol* 2004;5:233–6. [PubMed: 14985706]
4. Eckert DM, Kim PS. Design of potent inhibitors of HIV-1 entry from the gp41 N-peptide region. *Proc Natl Acad Sci U S A* 2001;98:11187–92. [PubMed: 11572974]
5. Weng Y, Yang Z, Weiss CD. Structure-function studies of the self-assembly domain of the human immunodeficiency virus type 1 transmembrane protein gp41. *J Virol* 2000;74:5368–72. [PubMed: 10799616]
6. Weissenhorn W, Calder LJ, Dessen A, Laue T, Skehel JJ, Wiley DC. Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 gp41 ectodomain expressed in *Escherichia coli*. *Proc Natl Acad Sci U S A* 1997;94:6065–9. [PubMed: 9177169]
7. Gabuzda DH, Lever A, Terwilliger E, Sodroski J. Effects of deletions in the cytoplasmic domain on biological functions of human immunodeficiency virus type 1 envelope glycoproteins. *J Virol* 1992;66:3306–15. [PubMed: 1583717]
8. Roux KH, Taylor KA. AIDS virus envelope spike structure. *Curr Opin Struct Biol* 2007;17:244–52. [PubMed: 17395457]

9. Lu M, Blacklow SC, Kim PS. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nat Struct Biol* 1995;2:1075–82. [PubMed: 8846219]
10. de Planque MR, Kruijtz JA, Liskamp RM, Marsh D, Greathouse DV, Koeppe RE 2nd, de Kruijff B, Killian JA. Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *J Biol Chem* 1999;274:20839–46. [PubMed: 10409625]
11. Muster T, Steindl F, Purtscher M, Trkola A, Klima A, Himmler G, Ruker F, Katinger H. A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J Virol* 1993;67:6642–7. [PubMed: 7692082]
12. Zwick MB, Labrijn AF, Wang M, Spenlehauer C, Saphire EO, Binley JM, Moore JP, Stiegler G, Katinger H, Burton DR, Parren PW. Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J Virol* 2001;75:10892–905. [PubMed: 11602729]
13. Zwick MB, Burton DR. HIV-1 neutralization: mechanisms and relevance to vaccine design. *Curr HIV Res* 2007;5:608–24. [PubMed: 18045117]
14. Dimitrov AS, Rawat SS, Jiang S, Blumenthal R. Role of the fusion peptide and membrane-proximal domain in HIV-1 envelope glycoprotein-mediated membrane fusion. *Biochemistry* 2003;42:14150–8. [PubMed: 14640682]
15. Munoz-Barroso I, Salzwedel K, Hunter E, Blumenthal R. Role of the membrane-proximal domain in the initial stages of human immunodeficiency virus type 1 envelope glycoprotein-mediated membrane fusion. *J Virol* 1999;73:6089–92. [PubMed: 10364363]
16. Resnick DA, Smith AD, Gesiler SC, Zhang A, Arnold E, Arnold GF. Chimeras from a human rhinovirus 14-human immunodeficiency virus type 1 (HIV-1) V3 loop seroprevalence library induce neutralizing responses against HIV-1. *J Virol* 1995;69:2406–11. [PubMed: 7884887]
17. Smith AD, Geisler SC, Chen AA, Resnick DA, Roy BM, Lewi PJ, Arnold E, Arnold GF. Human rhinovirus type 14:human immunodeficiency virus type 1 (HIV-1) V3 loop chimeras from a combinatorial library induce potent neutralizing antibody responses against HIV-1. *J Virol* 1998;72:651–9. [PubMed: 9420270]
18. Ding J, Smith AD, Geisler SC, Ma X, Arnold GF, Arnold E. Crystal structure of a human rhinovirus that displays part of the HIV-1 V3 loop and induces neutralizing antibodies against HIV-1. *Structure* 2002;10:999–1011. [PubMed: 12121655]
19. Ghiara JB, Ferguson DC, Satterthwait AC, Dyson HJ, Wilson IA. Structure-based design of a constrained peptide mimic of the HIV-1 V3 loop neutralization site. *J Mol Biol* 1997;266:31–9. [PubMed: 9054968]
20. Hagmann M. Computers aid vaccine design. *Science* 2000;290:80–2. [PubMed: 11183154]
21. Oomen CJ, Hoogerhout P, Bonvin AM, Kuipers B, Brugghe H, Timmermans H, Haseley SR, van Alphen L, Gros P. Immunogenicity of peptide-vaccine candidates predicted by molecular dynamics simulations. *J Mol Biol* 2003;328:1083–9. [PubMed: 12729743]
22. Nuzzaci M, Piazzolla G, Vitti A, Lapelosa M, Tortorella C, Stella I, Natilla A, Antonaci S, Piazzolla P. Cucumber mosaic virus as a presentation system for a double hepatitis C virus-derived epitope. *Arch Virol* 2007;152:915–28. [PubMed: 17238010]
23. Nair DT, Singh K, Siddiqui Z, Nayak BP, Rao KV, Salunke DM. Epitope recognition by diverse antibodies suggests conformational convergence in an antibody response. *J Immunol* 2002;168:2371–82. [PubMed: 11859128]
24. Ofek G, Tang M, Sambor A, Katinger H, Mascola JR, Wyatt R, Kwong PD. Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J Virol* 2004;78:10724–37. [PubMed: 15367639]
25. Ma B, Shatsky M, Wolfson HJ, Nussinov R. Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein Sci* 2002;11:184–97. [PubMed: 11790828]
26. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J Phys Chem B* 2001;105:6474–6487.
27. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 1996;118:11225–11236.

28. Gallicchio E, Levy RM. AGBNP: an analytic implicit solvent model suitable for molecular dynamics simulations and high-resolution modeling. *J Comput Chem* 2004;25:479–99. [PubMed: 14735568]
29. Felts AK, Harano Y, Gallicchio E, Levy RM. Free energy surfaces of beta-hairpin and alpha-helical peptides generated by replica exchange molecular dynamics with the AGBNP implicit solvent model. *Proteins* 2004;56:310–21. [PubMed: 15211514]
30. Ravindranathan KP, Gallicchio E, Friesner RA, McDermott AE, Levy RM. Conformational equilibrium of cytochrome P450 BM-3 complexed with N-palmitoylglycine: a replica exchange molecular dynamics study. *J Am Chem Soc* 2006;128:5786–91. [PubMed: 16637647]
31. Ravindranathan KP, Gallicchio E, Levy RM. Conformational equilibria and free energy profiles for the allosteric transition of the ribose-binding protein. *J Mol Biol* 2005;353:196–210. [PubMed: 16157349]
32. Felts AK, Paris K, Gallicchio E, Friesner RA, Levy RM. Predicting Long Loops with the AGBNP Implicit Solvent Model using Hierarchical Torsion Angle Sampling and Protein Local Optimization. *J Chem Theor Comp*. 2008in press
33. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters* 1999;314:141–151.
34. Tan K, Liu J, Wang J, Shen S, Lu M. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc Natl Acad Sci U S A* 1997;94:12303–8. [PubMed: 9356444]
35. Chan DC, Fass D, Berger JM, Kim PS. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 1997;89:263–73. [PubMed: 9108481]
36. Biron Z, Khare S, Samson AO, Hayek Y, Naider F, Anglister J. A monomeric 3(10)-helix is formed in water by a 13-residue peptide representing the neutralizing determinant of HIV-1 on gp41. *Biochemistry* 2002;41:12687–96. [PubMed: 12379111]
37. Barbato G, Bianchi E, Ingallinella P, Hurni WH, Miller MD, Ciliberto G, Cortese R, Bazzo R, Shiver JW, Pessi A. Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. *J Mol Biol* 2003;330:1101–15. [PubMed: 12860131]
38. Zwick MB, Jensen R, Church S, Wang M, Stiegler G, Kunert R, Katinger H, Burton DR. Anti-human immunodeficiency virus type 1 (HIV-1) antibodies 2F5 and 4E10 require surprisingly few crucial residues in the membrane-proximal external region of glycoprotein gp41 to neutralize HIV-1. *J Virol* 2005;79:1252–61. [PubMed: 15613352]
39. Ho J, Uger RA, Zwick MB, Luscher MA, Barber BH, MacDonald KS. Conformational constraints imposed on a pan-neutralizing HIV-1 antibody epitope result in increased antigenicity but not neutralizing response. *Vaccine* 2005;23:1559–73. [PubMed: 15694508]
40. Schibli DJ, Montelaro RC, Vogel HJ. The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles. *Biochemistry* 2001;40:9570–8. [PubMed: 11583156]
41. Ho J, MacDonald KS, Barber BH. Construction of recombinant targeting immunogens incorporating an HIV-1 neutralizing epitope into sites of differing conformational constraint. *Vaccine* 2002;20:1169–80. [PubMed: 11803079]
42. Cardoso RM, Zwick MB, Stanfield RL, Kunert R, Binley JM, Katinger H, Burton DR, Wilson IA. Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity* 2005;22:163–73. [PubMed: 15723805]
43. Salzwedel K, West JT, Hunter E. A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *J Virol* 1999;73:2469–80. [PubMed: 9971832]
44. Suarez T, Gallaher WR, Agirre A, Goni FM, Nieva JL. Membrane interface-interacting sequences within the ectodomain of the human immunodeficiency virus type 1 envelope glycoprotein: putative role during viral fusion. *J Virol* 2000;74:8038–47. [PubMed: 10933713]
45. Reithmeier RA. Characterization and modeling of membrane proteins using sequence analysis. *Curr Opin Struct Biol* 1995;5:491–500. [PubMed: 8528765]
46. McGaughey GB, Citron M, Danzeisen RC, Freidinger RM, Garsky VM, Hurni WM, Joyce JG, Liang X, Miller M, Shiver J, Bogusky MJ. HIV-1 vaccine development: constrained peptide immunogens show improved binding to the anti-HIV-1 gp41 MAbs. *Biochemistry* 2003;42:3214–23. [PubMed: 12641452]

47. Opalka D, Pessi A, Bianchi E, Ciliberto G, Schleif W, McElhaugh M, Danzeisen R, Geleziunas R, Miller M, Eckert DM, Bramhill D, Joyce J, Cook J, Magilton W, Shiver J, Emini E, Esser MT. Analysis of the HIV-1 gp41 specific immune response using a multiplexed antibody detection assay. *J Immunol Methods* 2004;287:49–65. [PubMed: 15099755]
48. Coeffier E, Clement JM, Cussac V, Khodaei-Boorane N, Jehanno M, Rojas M, Dridi A, Latour M, El Habib R, Barre-Sinoussi F, Hofnung M, Leclerc C. Antigenicity and immunogenicity of the HIV-1 gp41 epitope ELDKWA inserted into permissive sites of the MalE protein. *Vaccine* 2000;19:684–93. [PubMed: 11115689]
49. Kim M, Qiao Z, Yu J, Montefiori D, Reinherz EL. Immunogenicity of recombinant human immunodeficiency virus type 1-like particles expressing gp41 derivatives in a pre-fusion state. *Vaccine* 2007;25:5102–5114. [PubMed: 17055621]
50. Gallicchio E, Levy RM, Parashar M. Asynchronous replica exchange for molecular simulations. *J Comput Chem* 2007;29:788–794. [PubMed: 17876761]
51. Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, Felts AK, Halgren TA, Mainz DT, Maple JR, Murphy R, Philipp DM, Repasky MP, Zhang LY, Berne BJ, Friesner RA, Gallicchio E, Levy RM. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* 2005;26:1752–80. [PubMed: 16211539]
52. Qiu D, Shenkin PS, Hollinger FP, Still WC. The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii. *J Phys Chem A* 1997;101:3005–3014.
53. Bashford D, Case DA. Generalized born models of macromolecular solvation effects. *Annu Rev Phys Chem* 2000;51:129–52. [PubMed: 11031278]
54. Dmitrij Frishman PA. Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function, and Genetics* 1995;23:566–579.
55. Arnold E, Rossmann MG. Analysis of the structure of a common cold virus, human rhinovirus 14, refined at a resolution of 3.0 Å. *Journal of Molecular Biology* 1990;211:763–801. [PubMed: 2156077]
56. Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht HJ, Johnson JE, Kamer G, Luo M, Mosser AG, et al. Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* 1985;317:145–53. [PubMed: 2993920]
57. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 1993;26:283–291.
58. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 1983;221:709–713. [PubMed: 6879170]
59. Sridharan SA, Nicholls A, Honig B. A new vertex algorithm to calculate solvent accessible surface areas. *Biophys J* 1992;61:995a.
60. Peter S, Shenkin DQM. Cluster analysis of molecular conformations. *Journal of Computational Chemistry* 1994;15:899–916.

Figure 1. View of the representative modeled conformation of the NIm-II region of the cHRV1 HRV14:HIV-1 chimeric construct (A, left) and that of cHRV3 (B, right). The VP2 subunit and the inserted sequence are colored green and magenta respectively, including the side chains of the core ELDKWA epitope residues (DKW) and P158 (cHRV1 only) of the N-terminal linker. The VP1 subunit of HRV14 is shown in blue.

Figure 2.

Scatter plot of the C α RMSD values of the conformations of the ELDKWAS motif from the REMD ensembles at 313 K of the chimeric constructs cHRV1 (A), cHRV2 (B) and cHRV3 (C) with respect to the corresponding motifs of the peptide in the 1TJG crystal structure (x-axis) and the 1LCX NMR structure (y-axis).

Figure 3.

Scatter plot of the C α RMSD values of the conformations of the ELDKWS peptide in solution from the REMD ensembles at 313 K with respect to the corresponding sequence motifs of the peptide in the 1TJG crystal structure (x-axis) and the 1LCX NMR structure (y-axis). The conformations in the lower right panel are mostly α -helical, whereas the upper right panel holds mainly β -turn conformations.

Figure 4.

Structural comparison between the representative conformation of the α -helical conformational cluster of the ELDKWAS peptide (7-mer Peptide 1, cyan) and the ELDKWAS region of the 3₁₀ α -helical NMR structure (1LCX model 1, red) (left). Structural comparison between the representative conformation of the β -turn conformational cluster of the ELDKWAS peptide (7-mer Peptide 1, cyan) and the ELDKWAS motif of the crystal structure of the peptide complexed with the 2F5 Fab (1TJG, green) (right). The structures are superimposed with respect to the C α atoms of residues 662 to 668.

Figure 5. Scatter plot of the of the C α RMSD values of the ELDKWAS motif of the 21-mer LLELDKWANLWNWFDISNWLW peptide (A), and of 21-mer LLELDKWANLANAFDISNALA mutated peptide (B) with respect to the corresponding motifs of the peptide in the 1TJG crystal structure (x-axis) and the 1LCX NMR structure (y-axis).

Figure 6. Structural representation of a highly populated conformational state of the LLELDKWANLWNWFDISNWLW peptide in solution from the REMD ensemble at 310 K, illustrating the hydrophobic amino acids (red) involved in the formation of a tightly packed hydrophobic cluster.

Figure 7. Distribution of the solvent-accessible surface area (SASA) of the DKW core motif for the three chimeric constructs: cHRV1 (red), cHRV2 (brown), and cHRV3 (black). In cHRV3 the core of the epitope is significantly more solvent exposed than in the other two constructs.

Figure 8. Representation of hydrogen bonding network in the insertion region of the cHRV3 model. Hydrogen bonds are indicated by red dashed lines. The inserted peptide and the VP2 are drawn in magenta and green respectively. The residues involved in the hydrogen bond network are shown in stick representation.

Figure 9.

Representative structure of part of the cHRV3 chimeric virus from the REMD simulation. The ELDKWAS insert (magenta) is connected to the VP2 protomeric subunit. of cHRV3 (green). The VP1, VP3, and VP4 protomeric units are shown in blue, red, and gray, respectively. The Fab fragment of the 2F5 antibody is shown nearby to illustrate how it might bind to the epitope in the context of the virus. The heavy chain of 2F5 is shown in cyan and the light chain in orange.

Table 2

Degree of pairwise hydrophobic packing between pairs of tryptophan residues measured in the REMD simulation of the LLELDKWANLWNWFDISNWLW peptide among the conformations in which the structure of the DKW motif matched that of the corresponding residues in the complex with 2F5. Values closer to one indicate higher occurrence of hydrophobic interactions. Sequence numbering follows gp41 nomenclature.

	W670	W672	W678	W680
W666	0.72	0.77	0.22	0.46
W670		0.96	0.56	0.40
W672			0.64	0.29
W678				0.88