Hydrophobicity Maps of the N-Peptide Coiled Coil of HIV-1 gp41

Xavier Siebert‡ and Gerhard Hummer*

Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received October 12, 2001; Revised Manuscript Received December 26, 2001

ABSTRACT: Blocking HIV-1 viral entry into the host cell offers a promising new strategy for interfering with the HIV-1 life cycle. A major target of inhibitor design is to prevent binding of fusogenic gp41 C-peptides to the trimeric coiled coil of fusion-active N-peptides. Here, we map the hydrophobic character of the binding surface of the IQN17 peptide, a soluble analogue of the N-peptide coiled coil. The local binding affinity for a hydrophobic probe is determined by three methods: a hydrophobic force field, and molecular dynamics in solution analyzed by test particle insertion and inhomogeneous information theory. The regions of highest calculated hydrophobicity overlap with the positions of the hydrophobic anchor residues of the native C-peptides, and of two known inhibitors. Additional binding sites not exploited by these inhibitors are identified, and modifications for enhancing their binding affinity are suggested.

Infection by membrane-enclosed viruses is initiated by the fusion of the viral and cellular membranes, followed by the delivery of the viral genome into the cell cytoplasm (*1*). After activation of the virus, proteins from its glycoprotein coat undergo conformational changes which catalyze merging of viral and host cells, overcoming the repulsion between the two phospholipid bilayers. X-ray crystallography (*2*−*5*) and NMR (*6*, *7*) have produced an increasingly detailed view of the structures and structural changes involved in membrane fusion for viruses (for a recent review, see ref *8*), and similarly for the SNARE family of proteins (*9*).

In the case of HIV-1, the glycoprotein fragment gp41 is believed to anchor its hydrophobic N-terminal fusion peptides into the host cell membrane. This transiently exposes a trimeric N-peptide coiled coil to which the C-peptides of gp41 bind. The resulting six-helix bundle of N- and C-peptides is arranged as three hairpins. The prehairpin intermediate with exposed N-peptides offers an attractive target for drug development. Trapping the transient trimeric N-peptide coiled coil has been shown to block membrane fusion and thus viral entry (*10*−*17*), and the converse, trapping the C-peptides, has also been reported (*18*, *19*).

The C-peptides bind into three highly conserved symmetric hydrophobic grooves on the surface of the trimeric N-peptide coiled coil (*2*, *20*), leading to sensitivity of viral entry to mutations in that region (*21*−*25*). The grooves on the N-peptide coiled coil thus form a major target for gp41 inhibition. Indeed, it has been shown that peptides corresponding to segments of gp41 themselves inhibit viral entry into the host cell. In particular, peptides from the C-terminal region [e.g., DP-178 (*12*) or T649 (*26*)] bind to the N-terminal region and manifest inhibitory capacities, preventing the formation of the fusogenic gp41 hairpin core structure. Through mirror-image phage display, cyclic D-peptides have recently been identified that bind a soluble analogue (IQN17) of the N-peptide coiled coil of gp41 (*14*). Low-molecular weight compounds are being actively pursued (*16*, *17*, *27*, *28*).

The design of such small-molecule inhibitors greatly benefits from an experimental and theoretical characterization of the binding properties of the target molecule. While steric effects are well-understood, leading often to good correlations with binding affinities, more subtle interactions with entropic components have proved to be more elusive and are often not included in compound screening. Conformational flexibility in the binding site (*29*) and solvent-induced hydrophobic interactions (*30*, *31*) are two important yet poorly quantified factors determining binding affinities (*32*−*34*). If comparably large regions of a protein are involved in binding with relatively shallow binding pockets, as presented by the gp41 N-peptide coiled coil (*15*, *20*, *35*−*37*), including such entropic contributions is critical (*38*).

Several experimental methods have been developed for studying the local binding affinity for hydrophobic groups, in particular, X-ray crystallography and NMR with organic cosolvents (*39*−*43*) or under noble gas atmospheres (*44*). Theoretical methods for mapping protein surfaces often employ a combination of van der Waals interactions, electrostatics, and surface-area terms (*45*−*50*). The affinity of hydrophobic probes is directly related to fluctuations in the solvent structure at the protein surface (*51*, *52*). Here, we present three alternative approaches to mapping the binding affinity for a hydrophobic probe at the surface of the IQN17 peptide (*14*). Such affinity maps are useful both in the qualitative search for candidate molecules and in the quantitative screening with grid-based energy calculations. First, from molecular dynamics (MD) simulations of the peptide in solution, we construct a hydrophobicity map by test particle insertion (TPI). This map incorporates information about the mobility of the protein structure and the solvent. Second, we also use a generalization of the information theory (IT) for hydrophobic effects (*53*) to inhomo-

* To whom correspondence should be addressed.
‡ Present address: Department of Biophysics and Biophysical Chemistry, School of Medicine, Johns Hopkins University, Baltimore, MD 21205.

10.1021/bi0158526
Published on Web 02/07/2002

geneous systems as a computationally less demanding approach. Finally, to eliminate the high computational costs of molecular dynamics simulations, we calculate hydrophobicity maps from a force-field representation of hydrophobic effects (*54*), requiring only the static protein structure. The results are compared with three known structures: (1) the native six-helix N- and C-peptide coiled coil [PDB entry 1AIK (*20*)], (2) the N-peptide coiled coil of IQN17 in complex with a D-peptide inhibitor [PDB entry 1CZQ (*14*)], and (3) the native N-peptide coiled coil in complex with a nonpeptidic inhibitor [PDB entry 1FAV (*16*)]. We also discuss the relevance of the regions of high calculated hydrophobic affinity in light of some recent mutagenesis experiments identifying residues crucial for viral entry (*21*−*25*, *35*). On the basis of the hydrophobicity maps, we propose a simple modification of an existing peptidic inhibitor for enhancing its binding affinity.

MATERIALS AND METHODS

Molecular Dynamics Simulations. MD simulations were performed for a fragment of IQN17, the soluble analogue of the N-peptide coiled coil of gp41 (*14*). Since aggregation is not a concern in the MD simulations, we cleaved part of the solubilizing non-native C-terminus of the IQN17 peptide, and added acetyl and *N*-methyl blocking groups. The resulting sequence (Ac-EIARIKKLLQLTVWGIKQLQARIL-NMe) is identical to that of the native N-peptides in the underlined region of the hydrophobic groove, 558-AIEAN-NHLLQLTVWGIKQLQARIL-581 [numbering as in the precursor polypeptide gp160 (*20*)], which helps ensure that IQN17 correctly presents the target binding region. Indeed, a comparison of the crystal structures shows excellent agreement, with small differences confined to Leu-565 and Gln-577, and larger differences in the fully solvated Lys-574 (*14*, *20*). From the magnitude and location of these structural differences, we expect only a small influence on the binding properties.

The initial structure of the cleaved IQN17 was taken from its complex with a D-peptide inhibitor [PDB entry 1CZQ (*14*)] and used to build a trimeric coiled coil (24 amino acids in each strand). A crystallographic water molecule on the symmetry axis of the coiled coil was also included. The MD simulation of the coiled coil was performed by using the AMBER 6.0 program (University of California at San Francisco, San Francisco, CA) with the all-atom AMBER 94 force field (*55*) and TIP3P water molecules (*56*). All simulations were carried out under periodic boundary conditions. Long-range electrostatic interactions were treated by using the particle-mesh-Ewald method (*57*) with cubic spline interpolation, and a grid of 48^3 points with a width of ∼1 Å. A time step of 2 fs was used in the MD simulation, constraining the length of bonds involving hydrogen atoms by using the SHAKE algorithm (*58*). The pressure and temperature were kept at 1 bar and 300 K, respectively (*59*).

After energy minimization and MD (10 ps) in a vacuum, the peptide was solvated in an overall neutral rectangular box with a total of 2510 water molecules, 18 chloride ions, and six sodium ions (equivalent to an excess salt concentration of ∼130 mM). To equilibrate the solvent without distorting the protein, we artificially increased the mass of the protein atoms (10^3 times for H, 10^5 for the other atoms) and performed a 5 ps MD at 1000 K. Then we restored the regular masses and gradually heated the resulting structure from 10 to 300 K at a constant pressure of 1 bar to complete the equilibration. During the 4 ns production run, configurations were saved every 1 ps for analysis by TPI and inhomogeneous IT.

Test Particle Insertion. The affinity of a hydrophobic probe molecule at a point r near the surface of a protein is determined by the difference of the local and bulk-solvent excess chemical potentials [$\mu^{ex}(r)$ and μ^{ex}_{bulk}, respectively]. These excess chemical potentials are the free energies of transferring the probe molecule from the ideal gas phase to position r and to the bulk solvent, respectively. The chemical-potential difference defines a potential of mean force $W(r) = \mu^{ex}(r) - \mu^{ex}_{bulk}$. Here, we use the simplest hydrophobic probe molecule, a hard sphere excluding solvent and protein atoms. Attractive van der Waals interactions can be added by first-order perturbation theory (*60*). The chemical potentials are related to the probability p_0 of finding an empty volume v with a size and shape corresponding to the probe molecule [$\mu^{ex}(r) = -k_BT \ln p_0(r)$, where k_B is the Boltzmann constant and T the temperature].

The insertion probabilities $p_0(r)$ can be estimated perturbatively by TPI (*61*) from an equilibrium ensemble of structures without the need of additional free energy calculations. The structures of an equilibrium MD trajectory provide such an ensemble. Note that here these simulations are performed for IQN17 *without* a ligand bound. On a cubic grid with 0.3 Å spacing, p_0 is estimated by the fraction of successful insertions. For a given structure, a grid point belongs to a spherical cavity of radius R if all solvent and protein atoms i are at least a distance $R + R_i$ away. The atom radii R_i are 1.5 Å for N, 1.9 Å for C, 1.4 Å for O, 0 Å for H, 1.8 Å for Cl^-, and 0.85 Å for Na^+. As a probe radius we used 1 Å. Increasing the probe radius lowers the probability of successful insertions, and reduces the sampling efficiency.

The ensemble of structures sampled in the MD simulation is aligned to a common reference frame. We use rigid body rotations and translations to minimize the mean square distance from a reference structure (*62*) of a set of backbone atoms in the region of interest. In a recursive alignment, the crystal structure is used as an initial reference structure, and then successively replaced by the average of the aligned structures until convergence is achieved. The region of interest here is the hydrophobic pocket formed by 11 residues on each strand: Leu-565, Leu-566, Leu-568, Thr-569, Val-570, Trp-571, Gly-572, Ile-573, Lys-574, Leu-576, and Gln-577 (*20*). The local probabilities $p_0(r)$ are saved as crystallographic map files and analyzed with MidasPlus (*63*). These maps also exploit the symmetry of the molecule (three-helix coiled coil) to enhance the sampling. At each grid point, we consider the maximum value of p_0 among the three symmetry-related units, to identify all potential high-affinity regions.

Inhomogeneous Information Theory. In the IT model (*53*), p_0 is inferred by a maximum-entropy method that uses the mean m and variance σ^2 (and possibly higher moments) of the particle number fluctuations in a volume v with a size and shape corresponding to the probe molecule (here, spheres of radius R). For the solute radii considered here, the p_0 from

IT can be approximated by the Gaussian function $p_0 \approx (2\pi\sigma^2)^{-1} \exp(-m^2/2\sigma^2)$ (*53*, *64*). IT can easily be generalized to the inhomogeneous environment of a protein. However, the means and variances for observation volumes at points r then have to be determined from simulation data, unlike the bulk fluid case which permits a semianalytic treatment. The advantage of IT over TPI is that for larger probe radii, sampling the mean and variance is considerably more efficient than estimating p_0.

Hydrophobic Force Field. Hydrophobic interactions can be approximated by a force field with many-body terms (*54*). Here, we use a simplified implementation, dividing heavy protein atoms into polar (O and N) and nonpolar (C and S) atom classes. The potential of mean force of a methane probe near the rigid IQN17 coiled coil is calculated on a Cartesian grid. For the cavity terms in the hydrophobic force field, we sum all volume intersection terms (eq 3 of ref *54*). In addition to cavity terms, we include pair interactions $\omega_{ij}^{(2)}(r)$, approximated by those derived earlier for methane−methane pairs (*54*) in the case of nonpolar protein atoms, and by the analogously defined interaction of methane with water oxygen for polar atoms. Higher-order interactions $\omega^{(k\geq 3)}$ are not included (*54*). We use the proximity approximation (*65*, *66*), and include pair interactions of the probe only with the closest nonpolar and polar atom. In addition to the cavity and effective interaction terms, we add a first-order perturbation term for van der Waals interactions (eq 10 of ref *54*), again at the proximity level, and with water (*67*) and methane parameters (*68*) for polar and nonpolar atoms, with Lorentz−Berthelot combining rules. We note that upon binding this simplified treatment ignores the loss in van der Waals interactions of the probe with the solvent.

RESULTS AND DISCUSSION

The sequence and amino acid composition of the native C-peptides (*20*) and the D-peptide inhibitor differ (*14*). However, an alignment of the crystal structures of their complexes with the native N-peptides and the IQN17 peptide, respectively, reveals that they bind their target in similar modes, both exploiting the shallow hydrophobic binding pocket of the N-peptide coiled coil. Hydrophobic amino acid side chains that occupy related positions in the binding interface are underlined in the following sequences: C-peptide, 628-WMEWDREINNY-638; and D-peptide, 10-WAWLCAA-16. As discussed by Eckert et al. (*14*), D-Trp-10 overlaps almost perfectly with Trp-628. D-Trp-12 binds in the same region as Trp-631, but has its side chain in a perpendicular orientation. D-Leu-13 occupies a region between Trp-631 and Ile-635 of the N- and C-peptide complex. In the following, we will focus our attention on this pocket for which we expect the highest calculated hydrophobic affinities. We will first present the results of the molecular dynamics simulations, analyzed by two different methods: TPI and inhomogeneous IT. Then we compare those results with the computationally faster theoretical calculations based on a hydrophobic force field.

Molecular Dynamics Analyzed by Test Particle Insertion. Figure 1 shows an overview of the IQN17−peptide complex with the D-peptide inhibitor bound [PDB entry 1CZQ (*14*)]. Shown in green are the regions of highest hydrophobic affinity calculated from TPI at a contour level p_0 of 0.225

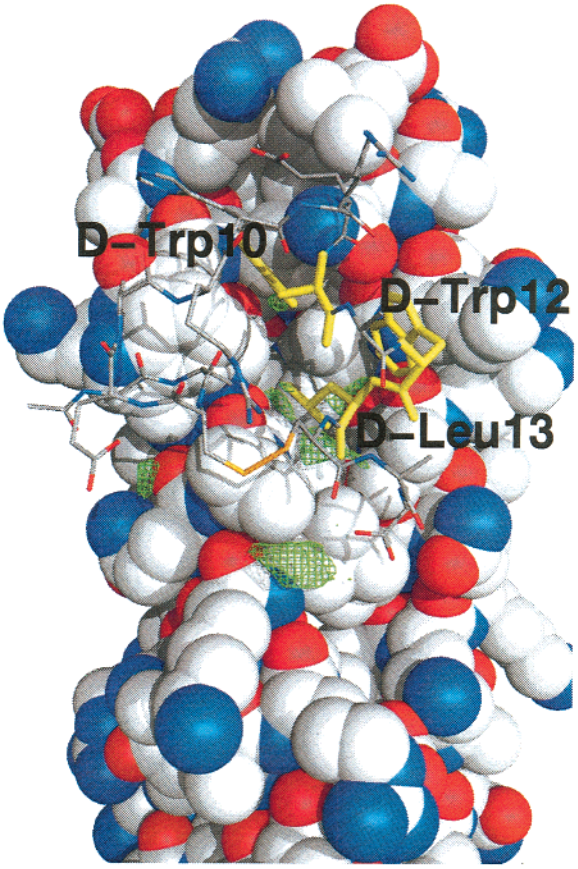

FIGURE 1: IQN17−peptide complex with the D-peptide inhibitor bound [PDB entry 1CZQ (*14*)]. The IQN17 and D-peptides are shown in CPK and solid-stick style, respectively (gray, carbon; red, oxygen; and blue, nitrogen). The residues forming the "hydrophobic anchor" of the D-peptide are shown in yellow. Shown in green are the regions of highest hydrophobic affinity calculated from TPI at a contour level of $p_0 = 0.225$ for a 1 Å probe.

(i.e., 22.5% success for inserting a 1 Å probe at a given location). At this contour level, we find five such regions within the hydrophobic pocket. Two high-affinity regions are outside, one near the C-terminus of the D-peptide and the other near Leu-566 and Leu-565 of another peptide strand. The hydrophobic affinity map is depicted in detail in Figure 2, which also shows the corresponding hydrophobic residues of the native C-peptides [PDB entry 1AIK (*20*)], as well as the ones from a nonpeptidic inhibitor [PDB entry 1FAV (*16*)] discussed below.

The hydrophobic residues of the C- and D-peptides occupying the binding pocket all fall within or near the regions of high calculated hydrophobic affinity. In particular, Trp-628 and D-Trp-10 contact the N-peptide with a high-affinity region between their indole rings and the N-peptide surface. Trp-631 overlaps with three of the calculated high-affinity regions. Interestingly, the indole nitrogen of Trp-631 also falls within a region of high calculated hydrophobicity. Indeed, in the D-peptide complex (*14*), that position is occupied by the δ_2 carbon of D-Leu-13. The δ_1 carbon of D-Leu-13 points toward the same high-hydrophobicity region as Ile-635.

Recently, Zhou et al. (*16*) determined the structure of an inhibitor bound to the HIV-1 gp41 trimeric core. A nonpeptidic moiety binds to the hydrophobic pocket in two different modes. One of the modes is not considered here because it

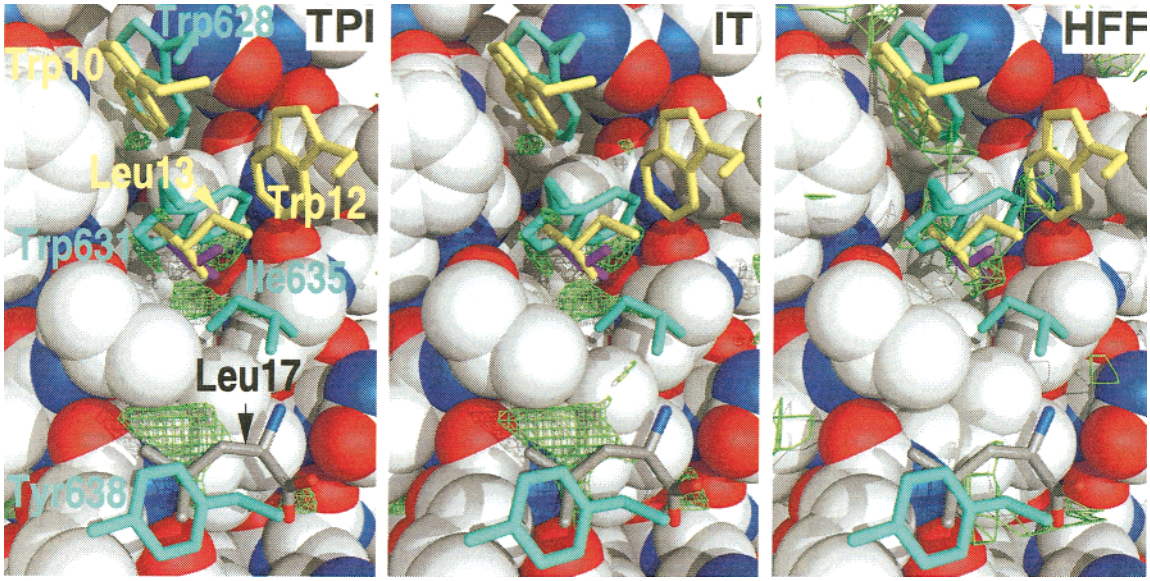

FIGURE 2: IQN17−peptide complex with the D-peptide inhibitor bound [PDB entry 1CZQ (*14*)]. Residues forming the hydrophobic anchor of the native C-peptides [PDB entry 1AIK (*20*)] and the D-peptide inhibitor [PDB entry 1CZQ (*14*)] are shown in light blue and yellow, respectively. The cyclopentyl group of a nonpeptidic inhibitor is shown in magenta [center of left panel; PDB entry 1FAV (*16*)]. Shown in green are the regions of highest hydrophobic affinity (left, TPI; middle, inhomogeneous IT, both at a contour level of $p_0 = 0.225$ for a 1 Å probe; and right, hydrophobic force field). Also shown near Tyr-638 is the residue D-Leu-17 (atom-colored sticks; with the carboxyl terminus) added to the C-terminus of the D-peptide inhibitor to enhance binding.

led to a 120° side chain rotation of Trp-571 in the N-peptide coiled coil, and would otherwise overlap with the protein matrix. In the second mode, however, the cyclopentyl group of the nonpeptidic inhibitor is adjacent to Trp-631 of the C-peptides, and D-Leu-13 of the D-peptide inhibitor in the respective complexes, and contacts three regions of high hydrophobic affinity, as shown in Figure 2.

From Figure 2, we expect that Trp-631 is the most tightly bound amino acid in the hydrophobic pocket, having extensive hydrophobic interactions at three sites of the indole ring. This observation is confirmed by a mutation study of C-peptide inhibition of viral entry. Chan et al. (*35*) find that mutation of Trp-631 to Ala dramatically worsens the ability of free C-peptides to inhibit viral entry, increasing the IC_{50} value ~30-fold compared to those of wild-type C-peptides. Less dramatic but still substantial effects were seen for mutations of Trp-628 and Ile-635 to Ala, with 5- and 2-fold increases in the IC_{50} values, respectively (*35*).

The importance of these regions of high calculated hydrophobic affinity for HIV-1 fusion is also emphasized by several mutagenesis studies of N-peptide residues (*21*−*25*). The N-peptide residue Ile-573 borders the regions of high hydrophobic affinity occupied by Trp-631 in the complex of N- and C-peptides (*20*). Mutations of Ile-573 to nonhydrophobic residues (Asp, Glu, Gly, and Ser) yield strongly suppressed fusogenic activity, whereas mutations to hydrophobic residues (Leu and Val) have only a small effect, with Ala producing an intermediate phenotype (*21*). The nearby Val-570 of the N-peptides also borders the region of highest calculated hydrophobic affinity occupied by Trp-631. Mutations of Val-570 indeed affect viral fusion (*23*−*25*). In particular, a mutation to Ala sharply lowers the stability of the model N34(L6)C28 of the six-helix bundle of N- and C-peptides, reducing its melting point by 14 °C (*25*). This is consistent with a sharp reduction in viral fusion activity caused by mutations of Val-570 to Ala (*24*, *25*), Asp, Glu, and Gly (*24*), and no effect on viral entry by a mutation to Ile.

In summary, the hydrophobicity maps from the MD calculations of the free N-peptide coiled coil model are supported by three independent crystal structures, one of the native C- and N-peptide complex (*20*) and two inhibitor complexes (*14*, *16*). Further support of the TPI calculations comes from mutation studies of the N-peptides (*21*, *23*−*25*), and a mutation study of the C-peptides which showed not only that the residues facing high calculated hydrophobicity regions are strongly affecting the inhibitory potency of free C-peptides but also that Trp-631 is critical for binding (*35*). This experimental observation correlates nicely with the extensive hydrophobic contacts of Trp-631 seen in the TPI hydrophobicity maps.

Outside the hydrophobic binding pocket already exploited by the D-peptide inhibitor, the TPI analysis of the MD trajectory identifies an additional high-affinity site. This more shallow site is located close to the C-terminal end of the D-peptide inhibitor, and is occupied in the wild-type complex of C- and N-peptides by the aromatic side chain of Tyr-638. This site is bordered by N-peptide residue Leu-565. A mutation of Leu-565 to Ala greatly destabilizes the model N34(L6)C28 of the six-helix bundle of N- and C-peptides and lowers its melting point by 20 °C (*25*). The L565A mutation also severely reduces membrane fusion activity (*25*). The large extent of this region of high hydrophobic affinity near Leu-565 (Figure 2) suggests a possible improvement in the design of the D-peptide inhibitor by addition of hydrophobic amino acids to its C-terminus. A molecular model of an elongated D-peptide inhibitor will be discussed below.

Molecular Dynamics Analyzed by Information Theory. For many large biomolecular systems, it is not feasible to perform the long MD simulations necessary to sample efficiently the local TPI cavity statistics. Instead of directly estimating the

insertion probability p_0, one can infer p_0 by using IT (*53*). In the simplest IT model, only the average and the variance of the local particle number fluctuations are needed as input. These can be estimated rather accurately even if the statistics of p_0 are poor, as in comparably short MD simulations. Figure 2 shows a hydrophobicity map calculated from IT at the same contour level and orientation as in Figure 1, again for a 1 Å probe. The IT and TPI contour surfaces agree well, thus validating the computationally more efficient IT method (*53*) even for the inhomogeneous environment near a protein surface.

Hydrophobic Force Field. In many practical applications, in particular, the screening of large compound databases, molecular dynamics in solution is prohibitively expensive. In such cases, one can attempt to approximate the hydrophobic affinity by statistical mechanical approaches, given the structure of the substrate protein, e.g., from X-ray or NMR measurements. Here, we use a hydrophobic force field that includes the significant many-body interactions while remaining computationally simple (*54*). We calculate the affinity for a methane probe on a Cartesian grid for the rigid PDB structure of the IQN17 peptide (*14*), thus neglecting protein flexibility. Figure 2 shows a hydrophobicity map calculated from the hydrophobic force field at a contour level chosen to give similar affinity as in the TPI and IT calculations for the 1 Å probe. Qualitatively, the results agree with those of TPI, despite the increased probe size, and the rigid structure. The most extensive region of high hydrophobicity is again located in the hydrophobic pocket, and overlaps with Trp-628, Trp-631, and Ile-635 of the C-peptide and D-Trp-10, D-Trp-12, and D-Leu-13 of the D-peptide. As in the TPI calculations, a more shallow yet strongly hydrophobic site is identified near Tyr-638. An additional weak site is found near the γ_2 carbon of Ile-635 and the β-carbon of D-Ala-16. That site is also present in the $p_0 = 0.225$ IT map (Figure 2), and in the TPI map at a lower contour level of $p_0 = 0.21$ (data not shown). Another small site near the α- and β-carbons of D-Ala-2, and in contact with the indole ring of Trp-571 of the N-peptides (top left in Figure 2), appears at $p_0 = 0.22$ in the TPI contour map.

Modeling of an Extended D-Peptide Inhibitor. On the basis of the hydrophobicity maps and the crystal structure of the N-peptide complex with C-peptides (*20*), we have modeled an extension of the D-peptide inhibitor by Eckert et al. (*14*). Specifically, we have added a D-leucine at the C-terminus to target the binding site predicted by the hydrophobicity maps, and occupied by Tyr-638 in the C-peptide complex. The newly added residue D-Leu-17 extends the short helix at the C-terminus of the D-peptide, with the added amide nitrogen of D-Leu-17 forming a bifurcated hydrogen bond to the carbonyl groups of D-Leu-13 and D-Cys-14. This backbone conformation allows the D-Leu-17 side chain to form tight hydrophobic interactions with Leu-565, Leu-568, and the aliphatic part of the Lys-28(IQN17) side chain (Figure 2). While the hydrophobic interaction with the lysine will be lost in the wild-type N-peptide coiled coil, with His-564 at the Lys-28(IQN17) position, a charged histidine (at low pH) could nevertheless form a salt bridge with the carboxylic end of the D-peptide. Salt bridges have been shown to be important for viral inhibition activity of compounds targeting gp41 (*69*).

CONCLUSIONS

The HIV-1 gp41 N-peptide coiled coil offers an attractive target for blocking viral entry into the host cell. During viral fusion, the native C-peptides of gp41 bind into three symmetric grooves on the surface of the transiently exposed N-peptide coiled coil. Within each groove, a prominent hydrophobic pocket, occupied by two tryptophan residues and one isoleucine residue of the C-peptide, has been the focus of drug development efforts (*10*, *14*−*17*, *26*−*28*, *37*, *70*). To design potent inhibitors that bind to this large yet shallow groove, it is important to optimize hydrophobic interactions, which play a major role in ligand binding and the formation of macromolecular complexes. Unlike van der Waals and electrostatic interactions, these hydrophobic interactions involve a large entropic component arising from the fluctuations of the macromolecule and solvent and are thus difficult to quantify.

Here, we use a multitier approach to calculate hydrophobicity maps of the surface of IQN17, a soluble analogue of the gp41 N-peptide coiled coil. (1) We perform molecular dynamics simulations, considering explicitly the motion of the target protein and solvent, and analyze them with (a) test particle insertion and (b) inhomogeneous information theory. (2) We employ a fast computational method, requiring only the static structure of the target molecule, by using a hydrophobic force field (*54*) that comprises the large entropic components of hydrophobic interactions and the resulting many-body effects.

The three calculations are in agreement with each other, and with the crystallographically determined binding modes of inhibitors (*14*, *16*) and C-peptide complexes (*20*), as well as mutation data for N-peptides (*21*, *23*−*25*) and inhibition of viral entry by C-peptides (*35*). The hydrophobic residues anchoring the inhibitors and C-peptides into the N-peptide pocket contact the regions of highest calculated hydrophobicity. Remarkably, our calculations not only consistently reproduce the hydrophobicity pattern inferred from these structures but also identify an additional hydrophobic binding site just outside the hydrophobic pocket. That site is exploited by Tyr-638 of the native C-peptides, but not by the inhibitors considered here. Its location near the C-terminus of the D-peptide inhibitor (*14*) suggests a possible improvement of inhibitor binding that can be achieved by slightly extending the length of the D-peptide used in the combinatorial search. As a simple model of such an extension, to be tested experimentally, we elongated the C-terminal helix by adding a D-Leu residue that occupies the additional hydrophobic binding site, while maintaining the structural integrity of the D-peptide.

Hydrophobic mapping of binding surfaces is proposed as a simple yet general tool for the design and optimization of ligands. However, quantitative applications of the hydrophobic force field, e.g., for calculating binding constants, will require extensive parametrization efforts. Compound databases can then be screened rapidly for binding to the gp41 N-peptides and other targets by precalculating the affinities for various probes on a grid. In the absence of such fully parameterized force fields, MD simulations can be combined with perturbation theory to construct affinity maps, thus complementing experimental (*39*−*44*) and other theoretical approaches (*45*−*50*), as has been done here for a

hydrophobic probe at the surface of the gp41 N-peptide coiled coil.

REFERENCES

1. Jahn, R., and Südhof, T. C. (1999) *Annu. Rev. Biochem. 68*, 863−911.
2. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) *Nature 387*, 426−430.
3. Fass, D., Harrison, S. C., and Kim, P. S. (1996) *Nat. Struct. Biol. 3*, 465−469.
4. Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) *Nature 371*, 37−43.
5. Malashkevich, V. N., Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9134−9139.
6. Caffrey, M., Cai, M. L., Kaufman, J., Stahl, S. J., Wingfield, P. T., Covell, D. G., Gronenborn, A. M., and Clore, G. M. (1998) *EMBO J. 17*, 4572−4584.
7. Koenig, B. W., Ferretti, J. A., and Gawrisch, K. (1999) *Biochemistry 38*, 6327−6334.
8. Eckert, D. M., and Kim, P. S. (2001) *Annu. Rev. Biochem. 70*, 777−810.
9. Sutton, R. B., Fasshauer, D., Jahn, R., and Brunger, A. T. (1998) *Nature 395*, 347−353.
10. Wild, C., Oas, T., McDanal, C., Bolognesi, D., and Matthews, T. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10537−10541.
11. Jiang, S., Lin, K., Strick, N., and Neurath, A. R. (1993) *Nature 365*, 113.
12. Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B., and Matthews, T. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9770−9774.
13. Chan, D. C., and Kim, P. S. (1998) *Cell 93*, 681−684.
14. Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A., and Kim, P. S. (1999) *Cell 99*, 103−115.
15. Sodroski, J. G. (1999) *Cell 99*, 243−246.
16. Zhou, G., Ferrer, M., Chopra, R., Kapoor, T. M., Strassmaier, T., Weissenhorn, W., Skehel, J. J., Oprian, D., Schreiber, S. L., Harrison, S. C., and Wiley, D. C. (2000) *Bio. Med. Chem. 8*, 2219−2228.
17. Jiang, S. B., and Debnath, A. K. (2000) *Biochem. Biophys. Res. Commun. 269*, 641−646.
18. Root, M. J., Kay, M. S., and Kim, P. S. (2001) *Science 291*, 884−888.
19. Louis, J. M., Bewley, C. A., and Clore, G. M. (2001) *J. Biol. Chem. 276*, 29485−29489.
20. Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) *Cell 89*, 263−273.
21. Dubay, J. W., Roberts, S. J., Brody, B., and Hunter, E. (1992) *J. Virol. 66*, 4748−4756.
22. Chen, S. S. L. (1994) *J. Virol. 68*, 2002−2010.
23. Weng, Y., and Weiss, C. D. (1998) *J. Virol. 72*, 9676−9682.
24. Weng, Y., Yang, Z., and Weiss, C. D. (2000) *J. Virol. 74*, 5368−5372.
25. Lu, M., Stoller, M. O., Wang, S., Liu, J., Fagan, M. B., and Nunberg, J. H. (2001) *J. Virol. 75*, 11146−11156.
26. Rimsky, L. T., Shugars, D. C., and Matthews, T. J. (1998) *J. Virol. 72*, 986−993.
27. Ferrer, M., Kapoor, T. M., Strassmaier, T., Weissenhorn, W., Skehel, J. J., Oprian, D., Schreiber, S. L., Wiley, D. C., and Harrison, S. C. (1999) *Nat. Struct. Biol. 6*, 953−960.
28. De Clercq, E. (2000) *Rev. Med. Virol. 10*, 255−277.
29. Carlson, M. L., and McCammon, A. J. (1999) *Mol. Pharmacol. 57*, 213−218.
30. Eisenhaber, F. (1999) *Perspect. Drug Discovery Des. 17*, 27−42.
31. Young, L., Jernigan, R. L., and Covell, D. G. (1994) *Protein Sci. 3*, 717−729.
32. Ringe, D. (1995) *Curr. Opin. Struct. Biol. 5*, 825−829.
33. Davis, A. M., and Teague, S. J. (1999) *Angew. Chem., Int. Ed. 38*, 736−749.
34. Gane, P. J., and Dean, P. M. (2000) *Curr. Opin. Struct. Biol. 10*, 401−404.
35. Chan, D. C., Chutkowski, C. T., and Kim, P. S. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 15613−15617.
36. Ji, H., Shu, W., Burling, F. T., Jiang, S. B., and Lu, M. (1999) *J. Virol. 73*, 8578−8586.
37. Moore, J. P., and Dragic, T. (1999) *Nature 401*, 759.
38. Cole, J. L., and Garsky, V. M. (2001) *Biochemistry 40*, 5633−5641.
39. Allen, K. N., Bellamacina, C. R., Ding, X. C., Jeffery, C. J., Mattos, C., Petsko, G. A., and Ringe, D. (1996) *J. Phys. Chem. 100*, 2605−2611.
40. Fitzpatrick, P. A., Steinmetz, A. C. U., Ringe, D., and Klibanov, A. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8653−8657.
41. Liepinsh, E., and Otting, G. (1997) *Nat. Biotechnol. 15*, 264−268.
42. Schmitke, J. L., Stern, L. J., and Klibanov, A. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 4250−4255.
43. Mattos, C., and Ringe, D. (1996) *Nat. Biotechnol. 14*, 595−599.
44. Otting, G., Liepinsh, E., Halle, B., and Frey, U. (1997) *Nat. Struct. Biol. 4*, 396−404.
45. Goodford, P. J. (1985) *J. Med. Chem. 28*, 849−857.
46. Majeux, N., Scarsi, M., Tenette-Souaille, C., and Caflisch, A. (2000) *Perspect. Drug Discovery Des. 20*, 145−169.
47. Scarsi, M., Majeux, N., and Caflisch, A. (1999) *Proteins: Struct., Funct., Genet. 37*, 565−575.
48. Heiden, W., Moeckel, G., and Brickmann, J. (1993) *J. Comput.-Aided Mol. Des. 7*, 503−514.
49. Bliznyuk, A. A., and Gready, J. E. (1999) *J. Comput. Chem. 20*, 983−988.
50. Gabdoulline, R. R., Wade, R. C., and Walther, D. (1999) *Trends Biochem. Sci. 24*, 285−287.
51. Cheng, Y. K., and Rossky, P. J. (1998) *Nature 392*, 696−699.
52. Carey, C., Cheng, Y.-K., and Rossky, P. J. (2000) *Chem. Phys. 258*, 415−425.
53. Hummer, G., Garde, S., García, A. E., Pohorille, A., and Pratt, L. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 8951−8955.
54. Hummer, G. (1999) *J. Am. Chem. Soc. 121*, 6299−6305.
55. Cornell, W. D., Cieplak, P., Bayley, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
56. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−935.
57. Darden, T., York, D., and Pedersen, L. (1993) *J. Chem. Phys. 98*, 10089−10092.
58. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327−341.
59. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Nola, A. D., and Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684−3690.
60. Garde, S., García, A. E., Pratt, L. R., and Hummer, G. (1999) *Biophys. Chem. 78*, 21−32.
61. Pohorille, A., and Pratt, L. R. (1990) *J. Am. Chem. Soc. 112*, 5066−5074.
62. McLachlan, A. D. (1979) *J. Mol. Biol. 128*, 49−79.
63. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
64. Garde, S., Hummer, G., García, A. E., Paulaitis, M. E., and Pratt, L. R. (1996) *Phys. Rev. Lett. 77*, 4966−4968.
65. Ashbaugh, H. S., and Paulaitis, M. E. (1996) *J. Phys. Chem. 100*, 1900−1913.
66. Garde, S., Hummer, G., García, A. E., Pratt, L. R., and Paulaitis, M. E. (1996) *Phys. Rev. E 53*, R4310−R4313.
67. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) in *Intermolecular Forces* (Pullman, B., Ed.) pp 331−342, Reidel, Dordrecht, The Netherlands.
68. Jorgensen, W. L., Madura, J. D., and Swenson, C. J. (1984) *J. Am. Chem. Soc. 106*, 6638−6646.
69. Jiang, S. B., and Debnath, A. K. (2000) *Biochem. Biophys. Res. Commun. 270*, 153−157.
70. D'Souza, M. P., Cairns, J. S., and Plaeger, S. F. (2000) *J. Am. Med. Assoc. 284*, 215−222.

BI0158526