Structural Models of the Bovine Papillomavirus E5 Protein

Toral Surti,[1] Ophir Klein,[2] Kathryn Aschheim,[1] Daniel DiMaio,[2] and Steven O. Smith[1]*
[1]*Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut*
[2]*Department of Genetics, Yale University, New Haven, Connecticut*

ABSTRACT **The bovine papillomavirus E5 protein is thought to be a type II integral membrane protein that exists as a disulfide-linked homodimer in transformed cells. Polarized-infrared measurements show that the E5 dimer in membrane bilayers is largely α-helical and has a transmembrane orientation. Computational searches of helix-helix conformations reveal two possible low-energy dimer structures. Correlation of these results with previous mutagenesis studies on the E5 protein suggests how the E5 dimer may serve as a molecular scaffold for dimerization and ligand-independent activation of the PDGF-β receptor. We propose that on each face of the E5 dimer a PDGF-β receptor molecule interacts directly with Gln17 from one E5 monomer and with Asp33 from the other E5 monomer. This model accounts for the requirement of Gln17 and Asp33 for complex formation and explains genetic results that dimerization of the E5 protein is essential for cell transformation. Proteins 33:601–612, 1998.** © 1998 Wiley-Liss, Inc.

Key words: platelet-derived growth factor receptor; E5 protein; bovine papillomavirus

INTRODUCTION

Viruses have evolved multiple strategies to transform host cells. One common route is to alter the expression of cellular genes by either integration of the viral genome into the cellular DNA or by interaction with nuclear proteins. Viruses can also encode membrane proteins that mimic cellular receptors or manipulate normal cell signaling.[1] Bovine papillomavirus (BPV) type 1 encodes a small integral membrane protein, the E5 protein, that appears to cause fibroblast transformation by activating endogenous growth factor receptors.[2] Remarkably, the E5 protein is the only BPV protein needed for cell transformation, even though it is only 44 residues in length.[2,3]

In transformed cells, the E5 protein is localized primarily to the membranes of the endoplasmic reticulum and Golgi apparatus,[4] and its carboxyl-terminal segment appears to extend into the lumen of these organelles. The E5 protein is able to interact with several different cellular proteins, including the epidermal growth factor receptor (EGFR), the human receptor for colony stimulating factor (CSF-1), the rat *neu* oncogene product $p185^{neu}$, the 16 kDa subunit of the vacuolar H^+ ATPase, and an α-adaptin-like molecule.[2] However, numerous studies indicate that the E5 protein causes fibroblast transformation by interacting with and activating the endogenous PDGF-β receptor.[5–11] The PDGF-β receptor is a receptor tyrosine kinase that is monomeric in unstimulated cells. Ligand-binding induces dimerization and activation of the receptor.[12] The E5 protein itself forms dimers in cell membranes and biochemical studies suggest that it induces ligand-independent dimerization of the receptor. Mutational analysis has identified specific residues in the transmembrane and juxtamembrane domains of both the E5 protein and the PDGF-β receptor that are required for complex formation and transformation, suggesting that these two proteins contact one another directly.[13–18] In contrast, the extracellular ligand binding domain of the receptor is not required for complex formation with the E5 protein, E5-induced receptor activation, or for mitogenic signaling.[6]

The N-terminal 32 amino acids of the E5 protein are largely hydrophobic and are thought to form a transmembrane α-helix.[2] Because aromatic residues are known to be favored in the polar regions of lipid bilayers,[19–21] and since Asp33 is the first charged residue in the sequence, the transmembrane segment of E5 is likely to be bounded by Trp5-Phe6 and Tyr31-Trp32. This segment of the E5 protein contains a single hydrophilic amino acid, Gln17. Many of the hydrophobic residues in the putative transmembrane sequence are conserved among E5 proteins from other fibropapillomaviruses. This suggests that the detailed packing of the hydrophobic side chains in the transmembrane helix interfaces

Abbreviations: ATR, attenuated total reflection; BPV, bovine papillomavirus; DMPC, dimyristoylphosphatidylcholine; EGFR, epidermal growth factor receptor; FSD, Fourier self-deconvolution; IR, infrared; PDGF, platelet-derived growth factor

Grant sponsor: National Institutes of Health; Grant numbers: GM 46732 and CA 37157.

Correspondence to: Steven O. Smith, Department of Biochemistry and Cell Biology, SUNY Stony Brook, 312 Life Sciences, Nicholls Road, Stony Brook, NY 11794. E-mail: steven.o.smith@sunysb.edu

Received 21 April 1998; Accepted 21 July 1998

may be important for stabilizing an active E5 dimer or E5-PDGF-β receptor complex.

Gln17 is identical among all of the fibropapillomavirus E5 proteins, and the nature of the amino acid at position 17 influences the efficiency of E5-induced cell transformation.[13–15,22–24] Several groups have varied the amino acid at this position and assessed the effects on transformation.[15,22,25] In general, hydrophobic substitutions at position 17 interfered with focus formation, whereas some charged and hydrophilic substitutions were tolerated. The results of Meyer et al.[15] suggested that the transformation efficiency of mutants correlated with the ability of the residue at position 17 to form interhelical hydrogen bonds. Recent studies systematically examining all possible position 17 mutants revealed a strong correlation between E5-PDGF-β receptor complex formation, receptor activation, and transformation for the various mutants in murine C127 cells.[42] Moreover, the identity of the amino acid at position 17 had marked effects on the extent of E5 dimerization, with hydrophobic substitutions in general allowing the least dimer formation.

Mutational analysis has also identified several highly-conserved residues in the hydrophilic C-terminal domain of the E5 protein that appear to be essential for cell transformation, including Cys37 and Cys39 which form the disulfide bonds involved in dimerization of the E5 protein.[13,14] The double mutation C37S/C39S abolishes transforming activity and also interferes with PDGF-β receptor phosphorylation and complex formation with the receptor.[13,14] While both Cys37 and Cys39 seem important in E5-induced transformation, their location is relatively flexible; moving cysteines to positions 34 and 42 or 36 and 40 did not impair transformation.[15] In addition, transforming activity was retained by mutants containing single cysteine residues at various positions in the carboxyl terminal segment. These results suggest that dimerization of the E5 protein is required for PDGF-β receptor activation. However, moving the two native cysteines to positions 35 and 41 eliminated the ability of the E5 protein to cause cell transformation.[15] This suggests that the exact position of the cysteine residues, and by inference the relative orientation of the E5 monomers in the dimer, is important for transformation. Asp33 in the C-terminal hydrophilic domain is also critical for E5-induced cell transformation. The single mutant D33V was unable to transform C127 cells, induce PDGF-β receptor phosphorylation or bind PDGF-β receptor.[13,14] Scanning mutagenesis confirmed that a negative charge in this domain of the E5 protein is critical for transformation of NIH3T3 cells.[15]

A series of mutant PDGF-β receptors has also been analyzed to map specific residues of the receptor that are important for complex formation with the E5 protein. A positive charge at position 499 in the extracellular juxtamembrane domain of the receptor was required for E5-PDGF-β receptor complex formation and E5-induced receptor activation, but not for activation by PDGF.[16] In addition, mutation of Thr513 to leucine in the transmembrane domain of the PDGF-β receptor inhibited the ability of the E5 protein to bind to and activate the receptor.[16] These results suggest that the amino acids at positions 499 and 513 in the PDGF-β receptor are important for complex formation. The E5 protein interacts with the PDGF β-receptor and not the closely related PDGF α-receptor.[11,26] Strikingly, both a positively charged residue around position 499 and the transmembrane threonine are absent from the α-receptor.

The PDGF-β receptor is a type I transmembrane protein with its N-terminal ligand binding domain extending away from the cytoplasm, and hence it has the opposite orientation as the E5 protein. Therefore, the transmembrane domains of the E5 protein and PDGF-β receptor are presumably aligned in an anti-parallel fashion. In this arrangement, Asp33 and Lys499 are located on the same side of the membrane and near the membrane interface. Furthermore, the spacing between Asp33 and Gln17 is similar to the spacing between Lys499 and Thr513. These considerations have previously suggested that Lys499 in the PDGF-β receptor may specifically interact with Asp33 in the E5 protein and that Thr513 may interact with Gln17.[15,16] This proposal is consistent with the opposite charges on Lys499 and Asp33 and with the tendency of polar groups, such as Gln17 and Thr513, to form hydrogen bonds when they are buried in hydrophobic membranes. However, although Lys499 and Thr513 in the PDGF-β receptor fall on the same face of an α-helix, Gln17 and Asp33 of the E5 protein are predicted to lie on opposite faces of an α-helix (Figure 1). Thus, it is not readily apparent how simultaneous Asp33-Lys499 and Gln17-Thr513 interactions could occur in the E5-PDGF-β receptor complex.

Despite its small size, there are no structural data on the E5 protein or on the interactions of the E5 protein with the PDGF-β receptor. The traditional high-resolution structural methods, X-ray diffraction and multidimensional NMR, generally encounter problems with hydrophobic membrane proteins. An alternative approach for defining the structure of the E5 dimer takes advantage of lower resolution methods and recent advances in computational search strategies.[27] The global secondary structure and orientation of membrane-spanning domains of integral membrane proteins can be assessed by polarized-infrared (IR) spectroscopy.[28] Establishing the secondary structure and orientation sets the stage for search strategies that have been developed for finding low-energy conformations of transmembrane α-helices. Brünger and coworkers have devised an algorithm that steps through different rotational orientations of two helices and calculates

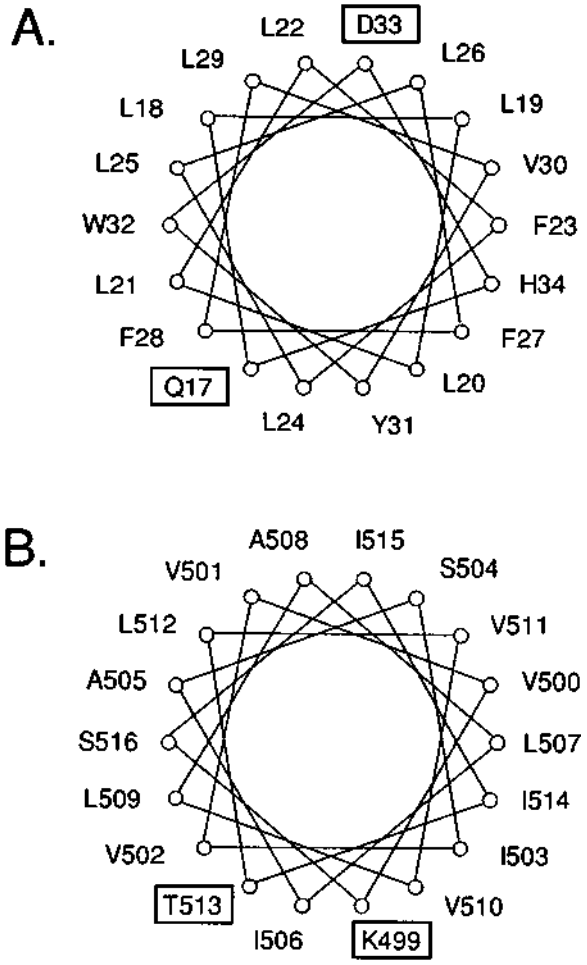

Fig. 1. Helical wheel diagrams of the E5 protein from Gln17 to His34 **(A)** and the transmembrane domain of the murine PDGF-β receptor from Lys499 to Ser516 **(B)**. The critical residues for productive interaction of the E5 protein with the PDGF-β receptor are shown in boxes. Gln17 and Asp33 lie on opposite faces of the E5 helix, while Lys499 and Thr513 lie on the same face.

the interaction energy of the resulting structures after molecular dynamics and energy minimization.[27]

In this study, we have used polarized-IR spectroscopy and computational methods to develop structural models of the E5 dimer. The polarized-IR results show that the E5 dimer is α-helical and spans membrane bilayers. Computational searches yield two low-energy dimeric structures. These structures are evaluated using several criteria derived from previous biochemical and mutational studies of the E5 protein and the E5-PDGF-β receptor complex, namely, (i) the dimer structure should be consistent with the transformation-sensitive residues in the bovine papillomavirus E5 protein, the conservation of residues among fibropapillomavirus E5 proteins, the effect of random hydrophobic insertions, and the existence of at least one intermolecular disulfide bond, (ii) the structure should account for the roles of Gln17 and Asp33 in E5-PDGF-β receptor complex formation, and should be consistent with a role for Thr513 and Lys499 of the PDGF-β receptor in complex formation, (iii) the structure should reflect the apparent role of Gln17 in E5 dimerization, and (iv) the structure should explain the correlation between E5 dimerization and transformation. The structures presented here are compatible with the criteria listed above and suggest a simple model by which complex formation results in receptor dimerization.

MATERIAL AND METHODS

Synthesis, Purification and Reconstitution of the E5 Protein

The 44-residue E5 protein was synthesized using solid-phase methods at the Keck Peptide Synthesis Facility at Yale University. Cysteine oxidation in the course of synthesis leads to the formation of dimers and higher-order oligomers. E5 dimers were HPLC purified on a TSK gel filtration column. The purified peptides were characterized by amino acid analysis and mass spectroscopy, and ran as a single dimer band using SDS PAGE. Reconstitution into oriented lipid bilayers was achieved by detergent dialysis. Lipid (dimyristoylphosphatidylcholine, DMPC), peptide (lyophilized), and detergent (octyl-β-glucoside) were dissolved in TFE. The TFE solution was lyophilized and the dry lipid/peptide/detergent mixture was rehydrated with phosphate buffer (10 mM phosphate and 50 mM NaCl, pH 7), such that the final concentration of octyl-β-glucoside was 5%. The rehydrated sample was stirred slowly for at least 6 h and then the octyl-β-glucoside was dialyzed away using Spectra-Por dialysis tubing (3000 MW cutoff) for 24 hours against phosphate buffer at a temperature above the lipid phase transition. The reconstituted membranes were sonicated briefly (30–60 s) in a ultrasonic bath and layered on a germanium crystal for IR measurements.

Polarized-IR Spectroscopy

Polarized-ATR-FTIR spectra of E5 were obtained on a Protege 440 IR spectrometer as described previously.[30] In general, 200 μL of multilamellar vesicle dispersions were spread on a 52 × 20 mm Ge internal-reflection element and dried using a flow of N_2 to form a uniform oriented multilamellar lipid-peptide film. Each sample spectrum represents the average of 1,000 scans acquired at a resolution of 4 cm^{-1}.

The measured dichroic ratio (R^{ATR}), defined as the ratio between absorption of parallel ($A_\parallel$) and perpendicular ($A_\perp$) polarized light, was used to calculate an order parameter S.

$$S = \tfrac{3}{2} \langle \cos^2\theta \rangle - \tfrac{1}{2}$$

$$S = 2\,(E_x^2 - R^{ATR}\,E_y^2 + E_z^2)/ \\ [(3\cos^2\alpha - 1)\,(E_x^2 - R^{ATR}\,E_y^2 - 2\,E_z^2)]$$

where θ is the angle between the helix director and the normal of the internal reflection element, and α is the angle between the helix director and the transition dipole moment of the amide I vibrational mode. The electric field amplitudes (E_x, E_y, and E_z) were calculated as in Arkin et al.[30] These equations are based on the assumption that the thickness of the deposited film (> 20 μm) is much larger than the penetration depth (~1 μm) of the evanescent wave.[41] A value of α = 42° was taken from recent studies on bacteriorhodopsin (S.C. Shekar and S.O. Smith, unpublished results). Lipid order parameters are obtained from the lipid methylene symmetric (2,852 cm^{-1}) and asymmetric (2,924 cm^{-1}) stretching modes using the same equation by setting α = 90°.

Computational Searches

The computational search strategy has been described previously by Adams et al.[27] Two canonical α-helices of residues 7–36 of BPV E5 were symmetrically rotated from 0° to 360°. At 20° increments, molecular dynamics (MD) simulations were performed using simulated annealing of all atomic coordinates. The structures were energy-minimized before and after the MD simulations. The parameters used for the MD simulations and energy minimization were the same as those used by Adams et al.[27] The starting geometries included both left-handed (+50°C) and right-handed (−50°C) crossing angles, the distance between the two helix axes was fixed at 10.5, 11.0, or 11.5 Å, and there was no translational offset between the helices. Five different molecular dynamics simulations were carried out for each starting geometry. The rotation and crossing angles were allowed to vary.

RESULTS AND DISCUSSION

E5 is Helical Transmembrane Dimer

Helical secondary structure is typical of the transmembrane domain of membrane proteins having a single hydrophobic membrane-spanning sequence.[29] This is likely to be the case in the PDGF-β receptor, in which the large hydrophilic extracellular and intracellular domains are found on opposite sides of the membrane. However, the situation is less clear for the E5 protein; it is only 44 residues in length, and it is possible to substitute the central hydrophilic glutamine with glutamic acid or lysine and retain transforming activity. As it is unusual to find charged amino acids in the membrane, these results raise the possibility that the E5 protein may be oriented parallel to the membrane so that the residue at position 17 is exposed to a polar environment. In order to establish the global secondary structure and orientation of the E5 protein, polarized-IR studies were undertaken on synthetic full-length dimeric E5 protein that had been reconstituted into model membrane bilayers formed from dimyristoylphosphatidylcholine (DMPC) (see Methods section). The frequency of the amide I vibration observed in IR spectra is sensitive to the secondary structure of the polypeptide. For α-helical secondary structure, the frequency typically ranges from ~1,650 to 1,660 cm^{-1}. In contrast, the amide I frequency for extended β-sheet structure is around 1,630 cm^{-1}.

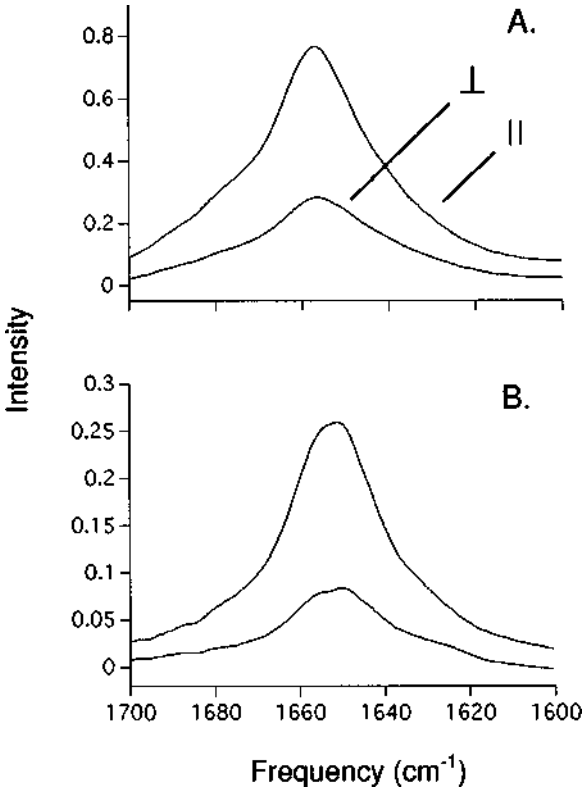

Fig. 2. Polarized-ATR-FTIR spectra of E5 in DMPC bilayers from H_2O **(A)** and D_2O **(B)**. Spectra were obtained with light polarized parallel (∥) or perpendicular (⊥) to the surface of the internal reflection element.

Figure 2 presents polarized attenuated total reflection (ATR) IR spectra of the E5 protein reconstituted into DMPC bilayers. Spectra obtained with parallel and perpendicular polarized light of oriented membranes layered from H_2O (Fig. 2A) exhibit a symmetric resonance at 1,657 cm^{-1}, characteristic of helical secondary structure. Fourier self-deconvolution (FSD) of the amide I band yields a single intense resonance at 1,657 cm^{-1}. There is no indication of β-structure at 1,630 cm^{-1}. It is possible to estimate the extent of helical secondary structure by comparing the intensity of the 1657 cm^{-1} band with the total intensity in the 1,600–1,700 cm^{-1} window and accounting for the residues (Asn, Gln) whose side-chain vibrations contribute to the amide I intensity.[30] This leads to an estimate that roughly 39 ± 3 residues in each E5 monomer are in α-helical secondary structure.

Since 18–25 hydrophobic residues are sufficient to span a membrane bilayer in an α-helical geometry,[29]

we next addressed whether any of the helical residues in the E5 dimer were exposed to water. Water accessibility can be determined by measuring the shift in frequency of the amide I vibration due to exchange of the amide proton for deuterium.[28,31] Those positions that are solvent accessible, i.e., not buried in the protein's hydrophobic core or in the membrane, exhibit amide I shifts of ~5–10 cm^{-1}. Spectra of the E5 dimer in oriented membranes layered from D_2O (Fig. 2B) exhibit two closely spaced resonances at 1,657 cm^{-1} and 1,650 cm^{-1}. The integrated intensities of the 1,657 cm^{-1} and 1,650 cm^{-1} bands obtained by Fourier self-deconvolution lead to the conclusion that ~30 ± 5% of the helical portion of the E5 protein is accessible to exchange with water. This conclusion is supported by the intensity drop observed in the amide II band at ~1,545 cm^{-1} in D_20 (data not shown). Together with the results above, this suggests that the helical hydrophobic transmembrane domain of E5 extends through the region of the phospholipid headgroups and that much of the hydrophilic C-terminal domain is helical.

Finally, the dichroic ratio of the 1,650–1,657 cm^{-1} band, defined as the ratio between absorption of parallel ($A_\parallel$) and perpendicular ($A_\perp$) polarized light, provides an indication of the orientation or tilt of the E5 protein relative to the membrane normal. The dichroic ratio of the E5 amide I band in H_2O is 3.2 (Fig. 2A). Water yields a broad weak band in the 1,600–1,700 cm^{-1} amide I window that is not dichroic and consequently lowers the measured dichroic ratio. This H_2O band is shifted out of the amide I window in D_2O. A qualitative comparison of the spectra in H_2O and D_2O shows that the amide I resonance narrows when the water background is removed. The dichroic ratio of the E5 amide I band in D_2O is 3.4, corresponding to a maximum helix tilt of ~20°. The actual helix orientation relative to the membrane normal is likely to be lower than 20° since any disorder in the orientation of the layered membrane bilayers or reconstituted E5 protein used for the polarized-IR measurements *lowers* the observed dichroic ratio.

Computational Searches Reveal Two Low-Energy Structures for the E5 Dimer

Several studies have shown that the E5 protein is dimeric in membranes,[3,15,22] leading to a model of the E5 protein as a dimer of two long α-helices. The polarized-IR results described above establish that the E5 protein spans cell membranes in α-helical secondary structure that extends into the extramembrane domain of the protein. Knowledge of the secondary structure and orientation of the E5 protein allows us to take advantage of computational search strategies that analyze interhelical interactions to predict the most probable structures of helix-helix dimers. These structures can subsequently be evaluated by comparison with biochemical and mutagenesis data. The success of combining global conformational searches, mutagenesis results, and structural data has been demonstrated by a series of studies that have established the dimer interface of glycophorin A.[32–34] Like the E5 protein, glycophorin A is a dimer with single membrane-spanning α-helices. The interacting residues in glycophorin A were first suggested by the ability of conservative amino acid substitutions in the transmembrane sequence to disrupt dimerization.[33,34] The same interhelical contacts were derived by Brünger and coworkers from conformational searches using the computational procedure described below.[27,32] Independently, specific interhelical distance constraints between interacting residues were established using magic-angle-spinning NMR that served to define the packing arrangement of several of the interacting side chains in lipid bilayers.[35] More recently, this model has been refined by distance contraints derived from high-resolution solution NMR studies of the glycophorin A transmembrane domain in detergent micelles.[36]

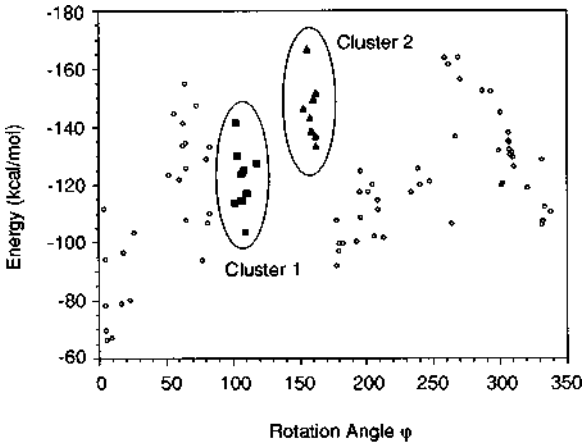

Fig. 3. Interaction energies of the individual structures in the computational search of wild-type E5 as a function of the rotation angle ϕ. The helix separation was held constant at 11.5 Å. Two clusters (circled) composed of at least ten structures with an rmsd of < 1 Å were found.

The computational search strategy was developed by Brünger and colleagues and has been described previously.[27] Low-energy conformations of helix dimers were identified by symmetrically rotating both helices through rotation angles ϕ1 and ϕ2 from 0–360° in 20° increments. For each starting geometry, five independent simulations were executed, each involving molecular dynamics (MD) and energy minimization. The rotation and crossing angles were allowed to vary, while the helix separation, measured from the helix axes, was held fixed at 10.5 Å, 11.0 Å, or 11.5 Å. This range of distances is consistent with interhelical separations in known membrane protein structures and allows interhelical

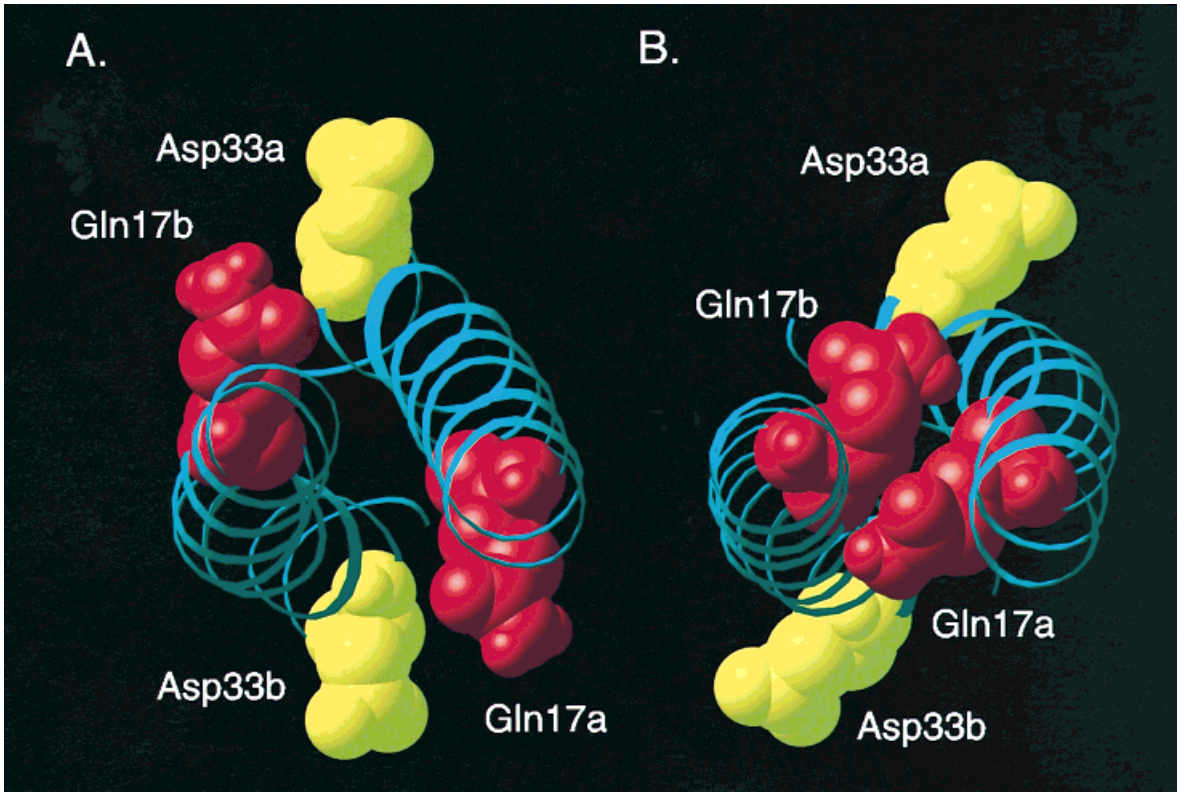

Fig. 4. Average molecular structures for cluster 1 **(A)** and cluster 2 **(B)** viewed down the E5 dimer axis showing the relative orientations of Gln17 (red) and Asp33 (yellow). Asp33 is oriented away from the dimer interface in both clusters. Gln17 is oriented away from the dimer interface in cluster 1, but packed in the interface in cluster 2. The helices in both dimers form left-handed coiled coils, and in both clusters, Asp33 and Gln17 of different E5 monomers lie on the same same face of the E5 dimer.

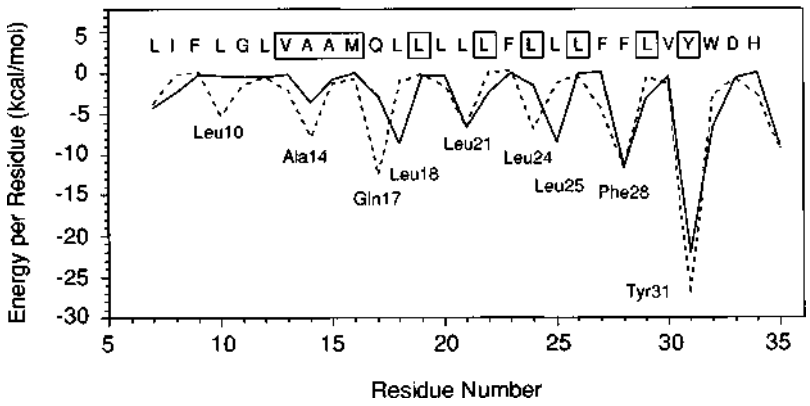

Fig. 5. Interaction energy per residue for the two average structures in cluster 1 (solid line) and cluster 2 (dotted line). The residues that are not conserved between E5 proteins from different species are boxed and largely fall outside of the dimer interface for both clusters. Hydrogen-bonding of the Gln17 side chain results in a significant drop in the interaction energy of this residue.

glutamine interactions (data not shown). The starting geometries included both left-handed and right-handed crossing angles. The sequence of the E5 protein used in the calculations extended from Leu7 to Glu36. This 30-residue segment contains the transmembrane sequence along with Asp33, but does not contain Cys37 or Cys39; it remains unclear whether the stretch from Cys37 to Phe44 is helical and whether both Cys37 and Cys39 are involved in disulfide bonds. Searches of the entire conformational space, as opposed to those of only symmetric structures, were carried out on the transmembrane domain alone, residues 7–30, and yielded comparable results (data not shown).

The results of one search are presented in Figure 3 which is a plot of the interaction energy for each

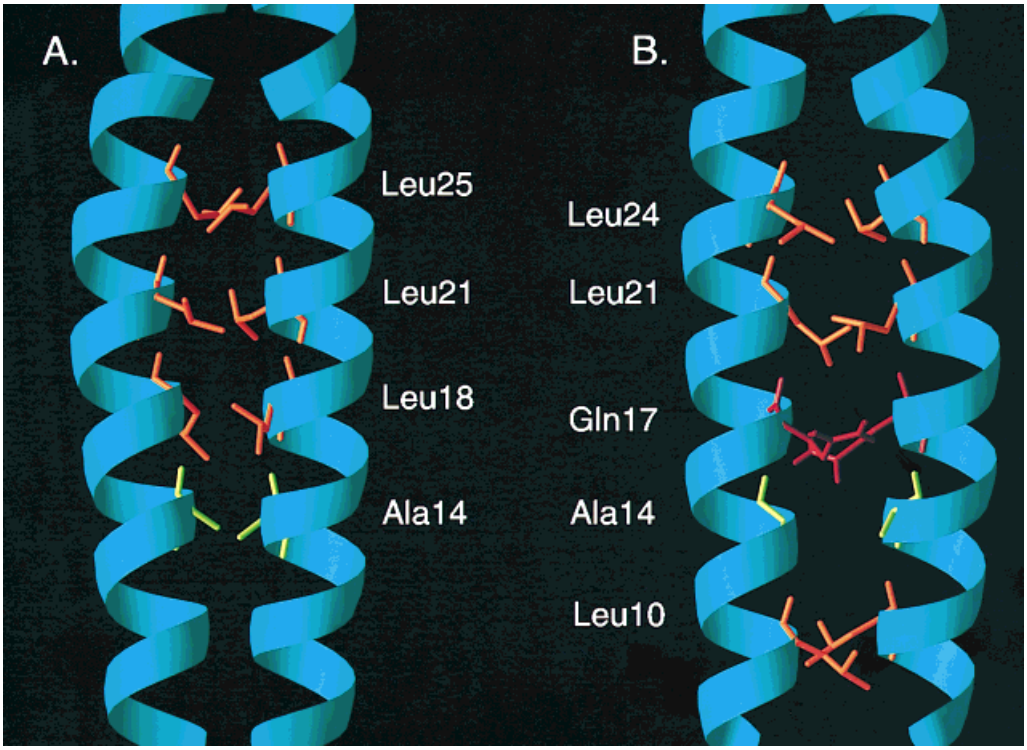

Fig. 6. Dimer interface of cluster 1 **(A)** and cluster 2 **(B)** illustrating the hydrophobic packing of leucine residues. The leucine side chains of one helix pack alongside the side chains of the opposing helix. The most notable difference in interfacial packing is that Leu18 is found in the interface of cluster 1 and Gln17 occurs in the interface of cluster 2.

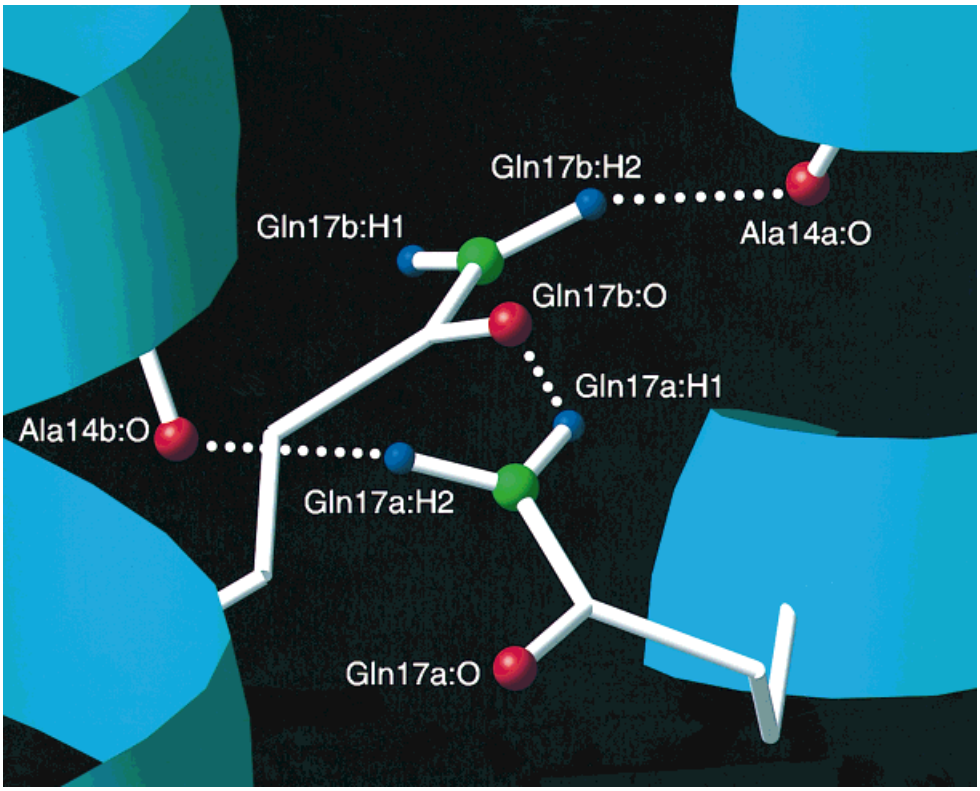

Fig. 7. Interhelical hydrogen-bonding of Gln17 in cluster 2. The lowest energy structure found in the computational searches resulted from the most favorable hydrogen-bonding interactions involving Gln17. The side-chain NH_2 groups of Gln17 are hydrogen-bonded across the dimer interface to the backbone carbonyls of Ala14. In addition, the NH_2 of Gln17 on helix A is hydrogen-bonded to the side chain C=O of Gln17 on helix B.

structure as a function of the average rotation angle ϕ. Initially, the rotation angles (ϕ1 and ϕ2) for both helices are equal and are separated from the next closest structure by 20°. During the cycle of molecular dynamics and minimization, the structures migrate from their initial geometries and often group together to form well-defined clusters of structures. A cluster is defined as a group of at least ten structures where the root-mean-square deviation of the atom positions is less than 1 Å between any given structure in the cluster and the next most similar structure. There are two clusters of structures, designated clusters 1 and 2, that satisfy this definition (Fig. 3). The structures in clusters 1 and 2 are roughly symmetric, i.e. the ϕ1 and ϕ2 angles are approximately the same for each structure in the cluster and vary over a range of ~10°. All of the structures in both clusters have a left-handed crossing angle of between 15.6° and 21.4°. Similar results were obtained in multiple searches. The group of structures at a rotation angle of ~300° did not satisfy the definition of a cluster and was not observed in searches with a helix separation of 10.5 Å and 11.0 Å. The cluster analysis allows one to perform a global search of conformational space in a relatively short period of time. The strategy is to generate a small pool of structures that can be evaluated on the basis of experimental data. Brünger and coworkers have argued that no simple criteria are capable of determining a correct structure based on computational approaches alone.[27] In particular, the use of the lowest energy as a single critierion can lead to the selection of incorrect models.[27] In the case of E5, the cluster analysis identified two robust local minima that are evaluated below.

Figure 4 presents the average molecular structures for clusters 1 and 2. Gln17 (red) and Asp33 (yellow) are highlighted to show their relative orientation. In cluster 1, the glutamine side chains are oriented away from the dimer interface, while in cluster 2 they are packed in the interface. In both clusters, Asp33 is oriented away from the interface, in a position where it can form electrostatic or hydrogen-bonding interactions with other proteins.

For clusters 1 and 2, the energetic contribution of each residue to the stability of the dimer is plotted in Figure 5. The interactions that contribute most significantly to the low energy of the helix dimer are distributed along the length of the interface (Figs. 5 and 6). A series of leucine residues (Leu10, Leu18, Leu21, Leu24, and Leu 25) along the interface pack in a leucine zipper-like motif, suggesting that these structures are largely stabilized by hydrophobic interactions. The packing of cluster 2 with Gln17 in the interface is similar to that of GCN4, the prototypical leucine-coiled coil where an Asn residue in the middle of the leucine-rich sequence hydrogen bonds across the dimer interface.[37,38] However, strikingly similar clusters to those described above resulted from calculations in which Gln17 was changed to leucine (data not shown), emphasizing the energetic contribution of hydrophobic packing in both models of E5 dimerization. In the calculations, the contribution of Gln17 to the interaction energy depends on the number and geometry of interhelical hydrogen bonds with significant stabilization occurring when the side-chain amide of Gln17 forms multiple hydrogen bonds across the interface. Although clusters were observed with Leu17 in the interface, they were less stable than with Gln17. This is consistent with experimental evidence that substitution of Gln17 with hydrophobic residues destabilizes, but does not completely disrupt, the E5 dimer.[42]

The residue with the largest energy contribution in both clusters is Tyr31. This tyrosine is well-packed in the interface with other aromatic residues and is hydrogen-bonded to Glu36 of the opposite helix. Moreover, a conservative Y31F substitution caused only a modest reduction in focus-forming activity [13]. However, the biological role of Tyr31 may differ from that suggested by interhelical hydrogen bonding in clusters 1 and 2. For example, the searches were done in the absence of lipids and consequently would not account for the possibility that this tyrosine may be involved in anchoring the helix in the membrane. Importantly, the same two clusters resulted from computational searches of the transmembrane domain alone (residues 7–30) emphasizing the point that stability of the E5 dimer results from interactions all along the interface and not from a single residue.

Sixteen of the twenty-six residues from Leu8 to Asp33 are identical in the E5 proteins of related fibropapillomaviruses (BPV and viruses isolated from deer, reindeer, and elk). Of the ten positions that are not identical among these viruses, eight of them (positions 13, 15, 16, 19, 22, 24, 26, and 29) are predicted to lie outside of the dimer interface in cluster 1, while seven of these eight are predicted to lie outside of the interface in cluster 2 (Figs. 5 and 6). Position 24 is in the interface in cluster 2. Of the three non-identical positions in cluster 2 that are proposed to lie in the interface, two are highly conserved in the sequences of related E5 proteins, namely, isoleucine at position 24 instead of leucine in BPV and tryptophan at position 31 instead of tyrosine. The only non-conserved position in the interface is Ala14, which is phenylalanine in the other three species. Searching the conformational space of an E5 mutant containing a phenylalanine-for-alanine substitution gave the same two clusters as described above for the wild type E5 dimer, showing that Phe14 could fit in the proposed helix-helix interface without disrupting the overall stability of the dimer (data not shown). The observation that the interfacial residues in clusters 1 and 2 largely coincide with the conserved residues provides strong

support for these clusters as structural models of the E5 dimer.

An absolutely conserved residue among the E5 proteins whose importance has been shown in several studies is Gln17. In general, structures with glutamine in the dimer interface, such as cluster 2, are of lower energy than structures with glutamine not in the interface, such as cluster 1 (Fig. 3). In cluster 2, Gln17 is in the interface and has the opportunity to form interhelical hydrogen bonds. Hydrogen-bonding of Gln17 across the dimer interface contributes to dimer stability. A comparison of interaction energies between simulations where the glutamine residues pack in the interface with different hydrogen bonding arrangements shows that the lowest-energy structures are those with the largest number of interhelical hydrogen bonds. Cluster 2 in Figure 4 has one interhelical hydrogen bond and yields a residue interaction energy of −12.5 kcal/mol. Figure 7 shows the result of a search that yielded the most favorable hydrogen-bonding interactions involving Gln17. The side chains of Gln17 are hydrogen-bonded across the dimer interface both to the backbone carbonyls of Ala14 and to each other. The three interhelical hydrogen bonds yield a residue interaction energy of −21.8 kcal/mol, and represents a significant contribution to dimer stability. Even in this conformation, there are two groups (Gln17B:NH and Gln17A:C=O) that are still free to hydrogen-bond, potentially to groups on the PDGF-β receptor.

To summarize, the polarized-IR data demonstrate that the E5 sequence is largely α-helical and oriented at an angle of ~20° or less to the membrane normal. Starting from the observation of helical transmembrane secondary structure, the computational searches found two symmetric and low-energy structures for the E5 dimer. As outlined in the introduction, four criteria have been used to evaluate these structural models. Both models are largely consistent with the conserved residues and transformation-sensitive residues (criterion i). It is more difficult to assess whether the structures are consistent with the random hydrophobic substitutions reported by Kulke et al.[23,24] since these mutants contained multiple substitutions from the wild-type sequence. Many of these substitutions were allowed if glutamine was retained in the sequence at position 17, emphasizing the crucial role of Gln17 in E5 activity. The sequences where hydrophobic substitutions seemed to be most detrimental involved substitution of phenylalanine for Leu24 and Leu25,[23] two residues predicted to be in the interface of the wild-type E5 dimer.

In both low-energy dimer structures, Gln17 and Asp33 are able to interact with the PDGF-β receptor (criterion ii). Asp33 is oriented away from the dimer interface in both clusters and has the potential to interact electrostatically with Lys499 on the PDGF-β receptor. The position of Gln17 is the most dramatic difference between the two structural models, since it is oriented towards the surrounding lipids in cluster 1, but is packed in the interface in cluster 2. Figure 7 shows that even in cluster 2, when Gln17 makes strong hydrogen bonds across the E5 dimer interface, there is still a free hydrogen bond donor and acceptor that can interact with Thr513 on the PDGF-β receptor. Therefore, it is not possible to rule out either structure based on the first two criteria. The observation that hydrogen bonding of Gln17 across the dimer interface stabilizes the E5 dimer is consistent with a role for Gln17 in dimerization (criterion iii) favoring cluster 2 as the structure of the E5 dimer. In particular, E5 dimer formation is reduced by replacing Gln17 with hydrophobic amino acids unable to form hydrogen bonds across the dimer interface (Meyer et al.,[15]; O. Klein et al.,[42]). Cluster 2 may also be favored energetically since the interaction energies calculated in Figure 3 do not reflect the unfavorable situation in cluster 1 where the polar glutamine side chains are exposed to a low dielectric environment and may not be in a position to form side-chain hydrogen bonds. However, since both clusters 1 and 2 are low-energy and differ only by a relative rotation of ~50°, it remains possible that both structures exist in membranes. It is also possible that binding of the PDGF-β receptor may stabilize one or the other of the dimer structures.

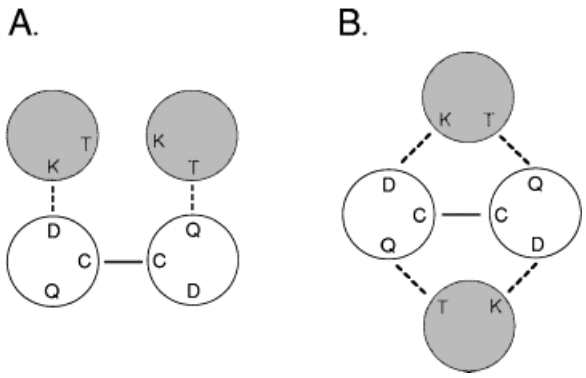

Fig. 8. Schematic of the E5-PDGF-β receptor complex viewed down the helix axis. The PDGF-β receptor (shaded circles) might interact with the E5 dimer (open circles) in either a *cis* arrangement **(A)** where the two PDGF-β monomers bind to the same face of the E5 dimer or in a *trans* arrangement **(B)** where they bind to opposite faces. In the *cis* geometry, each PDGF-β receptor monomer can interact with either Gln17 or Asp33, but not both simultaneously, since Gln17 and Asp33 lie on opposite faces of the E5 monomer. However, in the *trans* geometry, each PDGF-β receptor can interact with Gln17 on one E5 monomer and with Asp33 on the other.

The E5 Dimer May Serve as a Scaffold for Dimerization of the PDGF-β Receptor

Acceptable structural models of the E5 dimer and the E5/PDGF-β receptor complex must be consistent with the genetic requirements for complex forma-

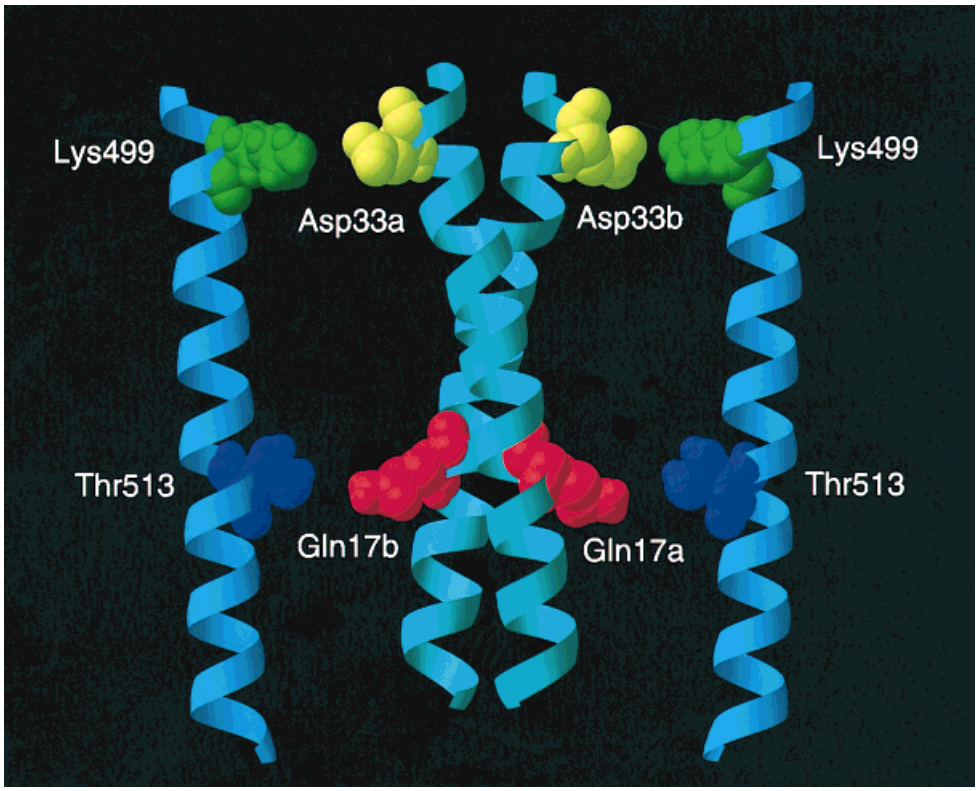

Fig. 9. E5 helix interface for docking of the PDGF-β receptor transmembrane domain in cluster 1. The PDGF-β receptor transmembrane helices are shown slightly displaced from the E5 dimer in order to highlight the complementary interactions between E5 and the receptor more clearly.

tion. Gln17 and Asp33 in the E5 protein are required for complex formation with the PDGF-β receptor, as are Lys499 and Thr513 of the PDGF-β receptor.[14,16] In addition, dimerization of the E5 protein appears essential for complex formation and fibroblast transformation (criterion iv), and the activated PDGF-β receptor in the complex appears dimeric. The discussion presented below is based on the assumption that the E5 protein and the PDGF-β receptor contact one another directly.

The overall geometry in the complex has important implications for the nature of the interactions between the component proteins. Figure 8 presents two possible arrangements of the E5 and PDGF-β receptor monomers in the E5-receptor complex. In the *cis* geometry, each PDGF-β receptor monomer interacts with either Gln17 or Asp33, but not both, since these two residues lie on opposite faces of the E5 monomer. However, in the *trans* geometry, the potential exists for a PDGF-β receptor molecule to interact with Gln17 and Asp33 on one face of the E5 dimer. If the transmembrane helices of the E5 protein and PDGF-β receptor assume an antiparallel orientation in the *trans* geometry, Lys499 of the receptor can interact with Asp33 of one E5 monomer while Thr513 of the same receptor molecule can interact with Gln17 of the other E5 monomer (Fig. 9). Since a second PDGF-β receptor molecule can undergo the same interactions on the opposite face of the E5 dimer, this arrangement can lead to receptor dimerization. Thus, the *trans* geometry overcomes the difficulties imposed by the position of the required Asp33 and Gln17 on opposite faces of each E5 monomer and can accommodate plausible interactions involving the amino acids known to be required for complex formation. Moreover, because the binding site for the PDGF-β receptor on the E5 protein is not present on the E5 monomer but is generated by E5 dimerization, this model can explain the requirement of E5 dimers for E5-induced cell transformation and provides a simple explanation for receptor dimerization and ligand-independent activation. Notably, both structural models of the E5 dimer allow docking of the PDGF-β receptor in a *trans* geometry.

The E5 dimer may represent a common motif of viral transformation mechanisms. The gp55 protein of the Friend spleen focus-forming virus, like the E5 protein of BPV, induces transformation by interacting with a cellular growth factor receptor. The gp55 protein, which is also thought to form a dimer in cell membranes, activates the erythropoetin (EPO) receptor.[39] Interactions involving transmembrane helices appear important for complex formation in this system as well.[40] It is possible that the gp55 protein

forms a scaffold for EPO receptor dimerization and thus causes constitutive activation of the receptor.

In this paper, polarized-IR and computational searches are combined with previous biochemical and mutagenesis data on the E5 protein to propose two models for the E5 dimer of BPV. These structures make specific predictions and can be further tested by biophysical and molecular biological experiments. Spectroscopic techniques, in combination with computational searches and mutational studies, may complement diffraction and NMR methods for determining membrane protein structure. The proposed structures appear able to explain the nature of the interactions between the E5 protein and its cellular target, the PDGF-β receptor. Furthermore, these studies illustrate how a small viral integral membrane protein can activate a growth factor receptor via interactions that are totally dissimilar from those that mediate activation by its normal ligand.

ACKNOWLEDGMENTS

We gratefully acknowledge Paul Adams and Axel Brünger for the programs and scripts used in the computational searches, and for many valuable discussions. O.K. was supported by an MSTP NIH Training Grant. The work was supported in part by grants from the NIH to S.S (GM 46732) and D.D. (CA 37157).

REFERENCES

1. Drummond-Barbosa, D., DiMaio, D. Virocrine transformation. Biochim. Biophys. Acta 1332:M1–17, 1997.
2. DiMaio, D., Petti, L., Hwang, E. The E5 transforming proteins of the papillomaviruses. Seminars in Virology 5:369–379, 1994.
3. Schlegel, R., Wade-Glass, M., Rabson, M.S., Yang, Y. The E5 transforming gene of bovine papillomavirus encodes a small hydrophobic protein. Science 233:464–467, 1986.
4. Burkhardt, A., Willingham, M., Gay, C., Jeang, K.T., Schlegel, R. The E5 oncoprotein of bovine papillomavirus is oriented asymmetrically in Golgi and plasma membranes. Virology 170:334–339, 1989.
5. Petti, L., DiMaio, D. Specific interaction between the bovine papillomavirus E5 transforming protein and the β-receptor for platelet-derived growth factor in stably transformed and acutely transfected cells. J. Virol. 68:3582–3592, 1994.
6. Drummond-Barbosa, D., Vaillancourt, R.R., Kazlauskas, A., DiMaio, D. Ligand-independent activation of the platelet-derived growth factor beta receptor: Requirements for bovine papillomavirus E5-induced mitogenic signaling. Mol. Cell Biol. 15:2570–2581, 1995.
7. Nilson, L.A., DiMaio, D. Platelet-derived growth factor receptor can mediate tumorigenic transformation by the bovine papillomavirus E5 protein. Mol. Cell Biol. 13:4137–4145, 1993.
8. Riese, D.J., DiMaio, D. An intact PDGF signaling pathway is required for efficient growth transformation of mouse C127 cells by the bovine papillomavirus E5 protein. Oncogene 10:1431–1439, 1995.
9. Petti, L., Nilson, L.A., DiMaio, D. Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein. EMBO J. 10:845–855, 1991.
10. Petti, L., DiMaio, D. Stable association between the bovine papillomavirus E5 transforming protein and activated platelet-derived growth factor receptor in transformed mouse cells. Proc. Natl. Acad. Sci. USA 89:6736–6740, 1992.
11. Goldstein, D.J., Li, W., Wang, L.M. et al. The bovine papillomavirus type 1 E5 transforming protein specifically binds and activates the beta-type receptor for the platelet-derived growth factor but not other related tyrosine kinase-containing receptors to induce cellular transformation. J. Virol. 68:4432–4441, 1994.
12. Heldin, C.-H. Dimerization of cell surface receptors in signal transduction. Cell 80:213–223, 1996.
13. Horwitz, B.H., Burkhardt, A.L., Schlegel, R., DiMaio, D. 44-amino-acid E5 transforming protein of bovine papillomavirus requires a hydrophobic core and specific carboxyl-terminal amino acids. Mol. Cell Biol. 8:4071–4078, 1988.
14. Nilson, L.A., Gottlieb, R.L., Polack, G.W., DiMaio, D. Mutational analysis of the interaction between the bovine papillomavirus E5 transforming protein and the endogenous beta receptor for platelet-derived growth factor in mouse C127 cells. J. Virol. 69:5869–5874, 1995.
15. Meyer, A.N., Xu, Y.F., Webster, M.K., Smith, A.E., Donoghue, D.J. Cellular transformation by a transmembrane peptide: Structural requirements for the bovine papillomavirus E5 oncoprotein. Proc. Natl. Acad. Sci. USA 91:4634–4638, 1994.
16. Petti, L.M., Reddy, V., Smith, S.O., DiMaio, D. Identification of amino acids in the transmembrane and juxtamembrane domains of the platelet-derived growth factor receptor required for productive interaction with the bovine papillomavirus E5 protein. J. Virol. 71:7318–7327, 1997.
17. Goldstein, D.J., Andresson, T., Sparkowski, J.J., Schlegel, R. The BPV-1 E5 protein, the 16 kDa membrane pore-forming protein and the PDGF receptor exist in a complex that is dependent on hydrophobic transmembrane interactions. EMBO J. 11:4851–4859, 1992.
18. Cohen, B.D., Goldstein, D.J., Rutledge, L., Vass, W.C., Lowy, D.R., Schlegel, R., Schiller, J.T. Transformation-specific interaction of the bovine papillomavirus E5 oncoprotein with the platelet-derived growth factor receptor transmembrane domain and the epidermal growth factor receptor cytoplasmic domain. J. Virol. 67:5303–5311, 1993.
19. Allen, J.P., Feher, G., Yeates, T.O., Komiya, H., Rees, D.C. Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The protein subunits. Proc. Natl. Acad. Sci. USA 84:6162–6166, 1987.
20. Haltia, T., Freire, E. Forces and factors that contribute to the structural stability of membrane proteins. Biochim. Biophys. Acta 1228:1–27, 1995.
21. Killian, J.A., Salemink, I., de Planque, M.R.R., Lindblom, G., Koeppe, R.E., Greathouse, D.V. Induction of nonbilayer structures in diacylphosphatidylinositol model membranes by transmembrane alpha-helical peptides: Importance of hydrophobic mismatch and proposed role of tryptophans. Biochemistry 35:1037–1045, 1996.
22. Sparkowski, J., Anders, J., Schlegel, R. Mutation of the bovine papillomavirus E5 oncoprotein at amino acid 17 generates both high- and low-transforming variants. J. Virol. 68:6120–6123, 1994.
23. Kulke, R., Horwitz, B.H., Zibello, T., DiMaio, D. The central hydrophobic domain of the bovine papillomavirus E5 transforming protein can be functionally replaced by many hydrophobic amino acid sequences containing a glutamine. J. Virol. 66: 505–511, 1992.
24. Horwitz, B.H., Weinstat, D.L., DiMaio, D. Transforming activity of a 16-amino-acid segment of the bovine papillomavirus E5 protein linked to random sequences of hydrophobic amino acids. J. Virol. 63: 4515–4519, 1989.
25. Sparkowski, J., Mense, M., Anders, J., Schlegel, R. E5 oncoprotein transmembrane mutants dissociate fibroblast transforming activity from 16-kilodalton protein binding and platelet-derived growth factor receptor binding and phosphorylation. J. Virol. 70: 2420–2430, 1996.
26. Staebler, A., Pierce, J.H., Brazinski, S., Heidaran, M.A., Li, W., Schlegel, R., Goldstein, D.J. Mutational analysis of the beta-type platelet-derived growth factor receptor defines the site of interaction with the bovine papillomavirus type 1 E5 transforming protein. J. Virol. 69: 6507–6517, 1995.

27. Adams, P.D., Engelman, D.M., Brunger, A.T. Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. Proteins 26:257–261, 1996.
28. Braiman, M.S., Rothschild, K.J. Fourier transform infrared techniques for probing membrane protein structure. Annu. Rev. Biophys. Biomol. Struct. 17:541–570, 1988.
29. Engelman, D.M., Steitz, T.A., Goldman, A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. 15: 321–353, 1986.
30. Arkin, I.T., Rothman, M., Ludlam, C.F. et al. Structural model of the phospholamban ion channel complex in phospholipid membranes. J. Mol. Biol. 248:824–834, 1995.
31. Tamm, L.K., Tatulian, S.A. Infrared spectroscopy of proteins and peptides in lipid bilayers. Q. Rev. Biophys. 30:365–429, 1997.
32. Treutlein, H.R., Lemmon, M.A., Engelman, D.M., Brunger, A.T. The glycophorin A transmembrane domain dimer: Sequence-specific propensity for a right-handed supercoil of helices. Biochemistry 31:12726–12732, 1992.
33. Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J., Engelman, D.M. Sequence specificity in the dimerization of transmembrane alpha-helices. Biochemistry 31:12719–12725, 1992.
34. Lemmon, M.A., Flanagan, J.M., Hunt, J.F. et al. A dimerization is driven by specific interactions between transmembrane alpha-helices. J. Biol. Chem. 267:7683–7689, 1992.
35. Smith, S.O., Bormann, B.J. Determination of helix-helix interactions in membranes by rotational resonance NMR. Proc. Natl. Acad. Sci. USA 92:488–491, 1995.
36. MacKenzie, K.R., Prestegard, J.H., Engelman, D.M. A transmembrane helix dimer: Structure and implications. Science 276:131–133, 1997.
37. Landschulz, W.H., Johnson, P.F., McKnight, S.L. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764, 1988.
38. O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:538–542, 1991.
39. Li, J., D'Andrea, A.D., Lodish, H.F., Baltimore, D. Activation of cell growth by binding of Friend spleen focus-forming virus gp55 glycoprotein to the erythropoietin receptor. Nature 343:762–764, 1990.
40. Zon, L.I., Moreau, J-F., Koo, J-W., Mathey-Prevot, B., D'Andrea, A.D. The erythropoietin receptor transmembrane region is necessary for activation by the Friend spleen focus-forming virus gp55 glycoprotein. Mol. Cell Biol. 12:2949–2957, 1992.
41. Harrick, N.J. "Internal Reflection Spectroscopy," New York: Interscience Publishers, 1967.
42. Klein, O., Polack, G.W., Surti, T., Kegler-Ebo, D.K., Smith, S.O., DiMaio, D. Role of glutamine 17 of the bovine papillomavirus E5 protein in PDGFβ receptor activation and cell transformation. J. Virology 72: in press, 1998.