Crystal Structure of the E2 Transactivation Domain of Human Papillomavirus Type 11 Bound to a Protein Interaction Inhibitor*

Yong Wang§, René Coulombe§, Dale R. Cameron%, Louise Thauvette§, Marie-Josée Massariol§, Lynn M. Amonz, Dominique Fink§, Steve Titolo§, Ewald Welchner§, Christiane Yoakim§, Jacques Archambault§, and Peter W. White§

From the §Department of Biological Sciences and ¶Department of Chemistry, Boehringer Ingelheim (Canada) Ltd., Laval, Quebec H7S 2G5, Canada and the ¶Department of Medicinal Chemistry, Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, Connecticut 06877

Received for publication, October 16, 2003, and in revised form, November 21, 2003
Published, JBC Papers in Press, November 22, 2003, DOI 10.1074/jbc.M311376200

Interaction between the E2 protein and E1 helicase of human papillomaviruses (HPVs) is essential for the initiation of viral DNA replication. We recently described a series of small molecules that bind to the N-terminal transactivation domain (TAD) of HPV type 11 E2 and inhibits its interaction with E1 in vitro and in cellular assays. Here we report the crystal structures of both the HPV11 TAD and of a complex between this domain and an inhibitor, at 2.5- and 2.4Å resolution, respectively. The HPV11 TAD structure is very similar to that of the analogous domain of HPV16. Inhibitor binding caused no significant alteration of the protein backbone, but movements of several amino acid side chains at the binding site, in particular those of Tyr-19, His-32, Leu-94, and Glu-100, resulted in the formation of a deep hydrophobic pocket that accommodates the indandione moiety of the inhibitor. Mutational analysis provides functional evidence for specific interactions between Tyr-19 and E1 and between His-32 and the inhibitor. A second inhibitor molecule is also present at the binding pocket. Although evidence is presented that this second molecule makes only weak interactions with the protein and is likely an artifact of crystallization, its presence defines additional regions of the binding pocket that could be exploited to design more potent inhibitors.

Human papillomaviruses (HPVs)1 are the etiological agents of malignant and benign lesions of the differentiating squamous or mucosal epithelium, notably of cervical cancer. Approximately 25 HPV types replicate in mucosal tissues of the anogenital tract. HPV16,-18, and -31 are the most prevalent “high-risk” types found in pre-cancerous or malignant lesions of the differentiating squamous epithelium, notably of cervical cancer. Approximately 25 HPV types replicate in mucosal tissues of the anogenital tract. HPV16,-18, and -31 are the most prevalent “high-risk” types found in pre-cancerous or malignant lesions of the cervix. HPV6 and -11 are the most common “low-risk” types, which cause benign genital warts (condyloma acuminata), a less serious condition but one of the most common sexually transmitted diseases (1). Currently, no specific antivirals are available for the treatment of HPV infections.

The small circular double-stranded DNA genome of papillomavirus is actively maintained as a multicopy episome in the nucleus of infected epithelial cells. This process is dependent on replication of the viral genome by the viral E1 and E2 proteins, in conjunction with the host DNA replication machinery. E2 is a sequence-specific DNA-binding protein that has a number of functions in the viral lifecycle. In addition to its role in the initiation of viral DNA replication, E2 is involved in regulating the transcription of viral genes (2–7), and in the segregation of the viral genome during cell division (8, 9). As a replication initiation factor, E2 binds with high affinity to specific sites located within the viral origin (ori) to help recruit it to the E1 helicase (10–13). Formation of a ternary complex between E1, E2, and the origin relies not only on the interaction of E1 and E2 with specific DNA sequences at the origin but is also critically dependent on a direct interaction between these two proteins (14–18).

The 40-kDa E2 protein exists in solution as a dimer and consists of an N-terminal transactivation domain (TAD) of ~200 residues, a C-terminal DNA binding/dimerization domain of ~100 residues, and a connecting “hinge” region of ~70 amino acids (19). The N-terminal TAD is the region of E2 that binds to E1 (14, 15, 18).

Crystal structures have been reported for the TAD of HPV16 (amino acids 1–201) (20) and for a truncated TAD of HPV18 (amino acids 66–215) lacking two N-terminal α-helices (21). The HPV16 and -18 TADs share ~45% sequence identity (see Fig. 1) and have a very similar L-shaped structure composed of a three-helix bundle, partly missing in the structure of the truncated HPV18 TAD, followed by an antiparallel β-sheet.

We recently described a series of small molecules that antagonize the E1-E2 protein-protein interaction by binding reversibly to the E2 TAD (22, 23). This class of inhibitors, termed “indandiones” because they feature an indandione system spirofused onto a substituted tetrahydrofuran ring, blocks assembly of the HPV11 E1-E2-ori ternary complex at submicromolar concentrations in vitro and abrogates viral DNA replication in a transient transfection assay. To our knowledge, these are the only small molecule inhibitors of HPV DNA replication with cellular activity.

Here, we report the three-dimensional structure of the HPV11 TAD and that of a TAD-indandione inhibitor complex. The HPV11 TAD structure is very similar to that of the analogous domains of HPV16 and -18. Inhibitor binding does not significantly affect the protein backbone but results in the
formation of a deep pocket that makes several well-defined interactions with the inhibitor. Residues that have previously been implicated in binding E1 or the indandione inhibitors, including Glu-12 (24), Glu-39 (25–27), and Glu-100 (23) are in or near this pocket.

**EXPERIMENTAL PROCEDURES**

**Proteins**—Wild-type (WT) HPV11 E2 TAD with an N-terminal poly-histidine tag (His-TAD, sequence MGHHHHHHHH followed by E2 amino acids 2–201) was purified from bacteria as described (23). To incorporate selenomethionine, the protein was expressed in *Escherichia coli* strain B834 (28), a methionine auxotroph, grown in DL30 medium (29) or near this pocket. – been implicated in binding E1 or the indandione inhibitors, rate selenomethionine, the protein was expressed in

201) was purified from bacteria as described (23). To incorpo-

– change into a buffer composed of 20 mM Tris (pH 8.0), 500 mM NaCl, 2 mM TCEP, 0.1 mM EDTA. Velocity data were collected at 60,000 rpm at 0.005-cm intervals, with 5 replicate readings per point. Protein concentrations were ~16 μM. Scans were spaced as closely as the absorbance optics would allow. Equilibrium data were collected successively at 26,000, 30,000, and 38,000 rpm using TAD concentrations of 2–8 μM. 5 replicate readings were taken at 0.003-cm intervals. Scans were taken every hour, and at each speed the system was judged to have reached equilibrium when no difference was observed in successive scans, as determined by subtraction using the Origin software (Microcal). Buffer density (1.01922 g/ml) and protein partial specific volumes at 20 °C (0.7292 and 0.7311 cm3/g for HPV11 His-TAD and HPV11 His-Lys-TAD, respectively) were calculated using the program SEDNTERP (30). Velocity data were analyzed using the program SVEDBERG 6.38 (31), and equilibrium data were analyzed using WinNonlin 1.06 (32). For equilibrium experiments, data were truncated at an absorbance of ~1.0, which corresponds to a TAD concentration of 7 μM. Very similar molecular weight estimates were obtained if more data, up to absorbance 1.2, was retained, but variances were higher and residuals were non-random.

**Isothermal Titrination Calorimetry**—Titration of inhibitor 1 into a solution of HPV11 His-Lys-TAD was performed using a VP-isothermal titration calorimetry instrument (MicroCal Inc.) in a buffer composed of 20 mM Tris (pH 7.5), 100 mM NaCl, 0.5 mM TCEP, 0.1 mM EDTA, and 1% (v/v) MeSO, at 20 °C, and was analyzed as described previously (23). Data from a blank run, using a solution of compound 1 injected into buffer without His-Lys-TAD, were subtracted from the inhibitor-protein run prior to fitting the data to a one-site binding model.

**Crytalization and Data Collection**—Final crystals of apo-HPV11 His-TAD containing either Met or Se-Met were obtained at 4 °C using micro-seeding in a hanging drop vapor diffusion plate. The precipitant solution was composed of 0.1 mM sodium succinate (pH 5.0), 18% PEG 5000 monomethyl ether, and 0.2 M (NH4)2SO4. Diffraction data on the Se-Met complex of His-Lys-TAD with inhibitor 1 were collected on an ADSC Q4 CCD at the National Synchrotron Light Source beamline X25 (Brookhaven National Laboratory, NY). Four datasets were collected from a single crystal, cooled to 100 K, at four different wavelengths near the selenium absorption edge (0.9789, 0.9794, 0.9743, and 0.9679 Å). The first dataset was selected for structure refinement; data collection statistics for this dataset are listed in Table I.

The complex of His-Lys-TAD with inhibitor 1 was formed by adding 1 μl of a 60 mM solution of 1 (BILH 434) in MeSO to 74 μl of His-Lys-TAD at 10 mg/ml in the precipitation buffer, and incubating the mixture on ice for 2–3 h. Crystals were then obtained by vapor diffusion, using a precipitant solution composed of 0.1 mM sodium citrate (pH 5.5) and 35% methyl-2,4-pentanediol. Diffraction data were collected on the National Synchrotron Light Source beamline X25 (Brookhaven National Laboratory, NY). Data from a single crystal cooled at 100 K were recorded on a Brandeis B4 detector (Brandeis University) mounted on a κ-axis goniometer (Etna-Nonius, The Netherlands). Data collection statistics are listed in Table I.
Structure Determination and Refinement—All data were processed with the HKL software Denzo and Scalepack (Table I). Initial phasing of the Se-Met HPV11 His-TAD Multim wavelength Anomalous Dispersion data was done using the software SHARP (33) with selenium sites located by SHELX (34). Model building into initial electron density maps was carried out using the software O (Alw6n Jones, Upsala University, Sweden) and model refinement was done with the software CNS (35). At a later stage, model building was also helped by molecular replacement phasing (CNX) using the published HPV16 coordinates (Protein Data Bank code IDTO (20)). 10% of the data was used as a test set for the Rfree calculation. The final model includes residues 3–121 and 128–196 (Table I), but 15 residues were modeled as alanines because of inadequate electron density for their side chains (Lys-109, Glu-17, Glu-48, Glu-107, Lys-120, Val-122, Leu-173, Ser-180, Thr-181, Asn-182, His-183, Ser-190, and Ser-195). The structure of the His-Lys-TAD1 complex was solved by molecular replacement using the refined HPV11 His-TAD model. This final model includes residues 2–196 (all side chains were modeled as the correct residue), two molecules of inhibitor, 27 water molecules, one molecule of Me2SO, and one partially defined buffer molecule (statistics in Table I). The refined atomic coordinates and structure factors have been deposited in the Protein Data Bank (accession codes 1R6K and 1R6N for the TAD and TAD:inhibitor 1 structures, respectively).

Computational Studies—The atoms making up the binding pocket were identified by applying the Site Finder application within the MOE software (Chemical Computing Group Inc., Montreal, Canada) to the crystal structure of the complex after both inhibitors, as well as solvent atoms, were removed. The probe radii were 1.4 and 1.8 A. The isolated donor acceptor radius was 3.0 A and the connection distance was 2.0 A.

Inhibitor-protein binding affinities were evaluated using the empirical scoring functions PMF (35), PLP (36), and LUDI (37). Each inhibitor-protein interaction was evaluated after removal of the other inhibitor as well as all solvent atoms (except one water molecule between His-29 and His-32 that appears to be hydrogen-bonded to both residues). For these calculations, hydrogen atoms were added to all residues and molecules.

E2-ori Complex Formation Assay—This assay was performed using in vitro translated E2 proteins as described previously (23). Briefly, binding reactions (80 ml) consisted of HPV11 E1-containing insect cell nuclear extract (titrated to give 50% of maximal activity with purified WT HPV11 E2, 2 ml of in vitro translated E2, and 0.4 ng of ori probe generated by PCR using [32P]dCTP. Negative controls contained E2 protein but no E1 extract. Signals were detected by scintillation counting using a TopCount NXT (PerkinElmer Life Sciences). Signals ranged from 30 to 4000 cpm for negative controls and 120 to 7000 for WT TAD (prior to blank subtraction). Binding data were analyzed by non-linear regression (Graphit 3.0, Erithacus Software Ltd.) using the quadratic equation,

\[ \text{Max} = \frac{(\text{TAD} + \text{E} + \text{Kd})}{4 \text{TAD} + 3 \text{E} + 3 \text{Kd}} \]  

where Max denotes the counts/min observed at saturation.

ATPase Assays—ATPase activity of purified recombinant E1 (40) was measured using a scintillation proximity assay as described previously (40, 41). The concentration of HPV11 E1 was 3 nm and ATP and magnesium acetate were used at 20 and 500 mM, respectively. After subtraction of background measured in reactions lacking E1, inhibition data were fit to a three-parameter logistic using SAS (release 6.12, SAS Institute Inc., Cary, NC). TAD mutants Y19A and E39A inhibited only weakly, such that their IC50 and I50 parameters could only be estimated.

RESULTS

Expression, Purification, and Characterization of the HPV11 E2 TAD—The HPV11 E2 TAD (amino acids 1–201, Fig. 1) was expressed in E. coli as a N-terminal polyhistidine-tagged protein. The purified protein eluted in a sharp peak when analyzed by analytical size-exclusion chromatography (Fig. 2A) and comparison of its elution time to those of molecular mass standards gave an estimated mass of 24.7 kDa, close to that expected for the monomer based on its sequence (23.8 kDa).

The HPV11 TAD also appeared to be monomeric by analytical ultracentrifugation. In an equilibrium sedimentation experiment, no significant trends in apparent mass were observed at increasing rotor speed or loading concentration (Table II). Such a trend would be observed if the protein formed a mixture of monomers and dimers under experimental conditions (42). Data from a centrifugation velocity experiment were best fit to an estimated average molecular mass of 22.5 kDa for a single species, based on individual parameters of S = 1.960 ± 0.003 Svedberg and D = 8.26 ± 0.03 Frick. Collectively, these results indicate that the HPV11 E2 TAD is fully monomeric in solution over the concentration range studied, up to ~20 lM.

An even more soluble version of the HPV11 TAD containing 4 lysines at its C terminus was also purified and was found to be monomeric by size-exclusion chromatography and analytical ultracentrifugation (data not shown). Binding of inhibitor 1 (Fig. 3) to this Lys-tailed TAD was analyzed by isothermal titration calorimetry. Formation of a 1:1 complex was observed with a Kd value of 40 nM (Fig. 2C).

Crystal Structure of the Apo-HPV11 E2 TAD—The structure of the HPV11 TAD was determined from Multidwavelength Anomalous Dispersion diffraction data obtained using crystals of the protein labeled with Se-Met (see “Experimental Procedures”). Overall the structure of the HPV11 TAD shown in Fig. 4 is very similar to that previously reported for the analogous domain of HPV16. The all-Co superposition of the HPV11 and -16 structures (187 residues) yielded an r.m.s. deviation of 1.54 A, and all secondary structural elements are conserved. The major differences are a slight twist (~1°) in the relative orientation of the a-helical and b-sheet subdomains and to variations in the position of loops, in particular those comprised of residues 140–143 and 160–165, which are involved in crystal packing in the HPV16 and HPV11 structures, respectively. The N-terminal (amino acids 3–94) and C-terminal (amino acids 95–195) subdomains can be individually superimposed with r.m.s. deviations of 0.85 and 1.14 A, respectively. Interestingly, the HPV11 protein crystal lattice was distinct from that reported for HPV16, such that a different monomer-monomer interface was observed in the HPV11 crystal structure.

Crystal Structure of an HPV11 E2 TAD-Inhibitor Complex—We obtained crystals of the more soluble Lys-tailed TAD in complex with compound 1, albeit using different conditions than those used to crystallize the apo-His-TAD. The space
group of these crystals was also different from any of the previous TAD structures (Table I). The conformation of the HPV11 TAD within the complex (Fig. 5A) is very similar to that observed for the apoprotein; all backbone Cα could be superimposed with an r.m.s. deviation of 0.88 Å (187 residues). The two inhibitor molecules are both in the free carboxylate form (as drawn in Fig. 3), one of two distinct forms that this class of inhibitors adopts in solution (43). The presence in the crystal
structure of the free carboxylate form confirms a prediction that this form is the one with inhibitory activity (43).

Because we had demonstrated that the TAD and inhibitor 1 form a 1:1 complex in solution, we were surprised to find two molecules of inhibitor bound per protein monomer (Fig. 5A). We believe that the molecule referred to below as inhibitor A is the one that binds tightly in biological assays (see below). In contrast, inhibitor B appears to interact only weakly with a secondary binding pocket, formed not only by the TAD protein alone, but also in part by one surface of inhibitor A, as well as by residues from an adjacent protein monomer in the crystal lattice. In fact, inhibitor B has approximately equal surface area in contact with each of the two TAD molecules, as well as with inhibitor A (data not shown). The close contact between the two inhibitor molecules may distort the observed interaction between the TAD and the dichlorophenyl ring of inhibitor A (see below). The TAD-TAD interface that incorporates inhibitor B is distinct from that in the apo structure, and in fact does not result in formation of a dimer, but rather an infinite array (along the 41 screw axis of the crystal), which is unlikely to exist in solution except at very high protein and inhibitor concentrations.

To further demonstrate that the region of the TAD near inhibitor A constitutes a better binding pocket than does the one near B, a computational analysis of the protein surface was performed using the Site Finder application within the MOE software package (see “Experimental Procedures”). This program uses an α shape-based algorithm (44) to identify concave regions on the protein surface. Furthermore, it evaluates whether these cavities have an appropriate composition of hydrophobic and hydrophilic atoms, as hydrophilic cavities are likely to be highly solvated and thus interact only weakly with organic molecules in aqueous solution. The TAD surface was evaluated after removal of the two inhibitors, and residues that comprised the best small molecule-binding pocket are highlighted in Fig. 5B. Inhibitor A makes a number of close contacts with these highlighted residues, as described below. Inhibitor B is located on the periphery of this pocket and does not make any well defined hydrogen bonding or Van der Waals interactions. The dichlorophenyl ring of inhibitor B is ~4 Å from the nearest pocket atom; other portions of the molecule are further from the protein surface (indandione moiety) or make closest contact with non-pocket atoms (thiadiazolylphenyl moiety).

In a separate computational study, binding energies for each inhibitor molecule to the TAD (in the absence of the other) were predicted using three different empirical protein-binding scoring functions (Table III). As expected, results significantly favored binding of inhibitor A relative to B, with predicted binding energies 2-fold higher for inhibitor A. Taken together, these different methods suggest that inhibitor A is solely responsible for binding to the TAD.

**Table II**

Analytical ultracentrifugation data for HPV11 His-TAD

| Speed  | [TAD]a | c | M^b | s^c |
|--------|--------|---|-----|-----|
| rpm    | μM     | cm^-2 | Da  | ODU |
| 26,000 | 8      | 1.76 | 22,500 | 0.008 |
| 26,000 | 4      | 1.85 | 23,600 | 0.003 |
| 26,000 | 2      | 1.83 | 23,500 | 0.004 |
| 30,000 | 8      | 2.45 | 23,600 | 0.008 |
| 30,000 | 4      | 2.47 | 23,800 | 0.005 |
| 30,000 | 2      | 2.43 | 23,300 | 0.003 |
| 38,000 | 8      | 3.94 | 23,600 | 0.007 |
| 38,000 | 4      | 3.93 | 23,600 | 0.009 |
| 38,000 | 2      | 3.84 | 23,000 | 0.004 |

a Loading concentration, prior to run.

b Reduced molecular weight (49).

c Molecular weight calculated from σ assuming partial specific volume for His-TAD of 0.7292 cm^3/g and a solution density ρ = 1.01922.

d Square root of the variance of the fit.
Fig. 5. Structure of the HPV11 TAD-inhibitor 1 complex. A, stereo view of the complex of two inhibitor molecules bound to the HPV11 TAD (1R6N). Inhibitor molecule B is above molecule A. B, stereo view of the molecular surface of HPV11 TAD with inhibitor molecules A (bottom) and B (top). Spheres corresponding to locations within the binding site where ligand atoms are predicted to make particularly favorable interactions with the protein are shown in yellow and amino acids that comprise this small molecule binding pocket are shown in magenta. The indandione moiety of inhibitor B (left-most portion in the view shown) is well above the protein surface, ~7 Å from the nearest residue in the binding pocket. C, stereo view of selected amino acids within 4 Å of inhibitor A are shown as they appear in the apo structure (magenta) and in the inhibitor-TAD complex (green). Inhibitor A is shown in orange. The relative positions of residues are based on an all C-α superposition of the two structures. Thin green lines indicate hydrogen bonds between the inhibitor carboxylate and main chain amide nitrogens of residues 99 and 100.
for inhibition at the low concentrations used in biological assays and that the presence of inhibitor B might be an artifact of crystallization. This suggestion is further supported by the mutational analysis of the inhibitor binding pocket, presented below. Although an artifact, it is possible that the fortuitous presence of this second molecule aided in crystallization of the complex.

Inhibitor A lies within 4.5 Å of 12 residues (marked by asterisks in Fig. 1), and the backbone superposition for these in the apo and complex structures gives an r.m.s. deviation of only 0.28 Å. However, binding of this inhibitor induces significant movement of several amino acid side chains. The indandione moiety of the inhibitor is buried in a pocket that is not present in the apo structure, but results from rotations of the side chains of Tyr-19 (χ₁, 85°; χ₂, 40°) and His-32 (χ₂, 90°) (Fig. 5C).

The new positions of Tyr-19 and His-32 not only open up the pocket, but also allow the indandione moiety of the inhibitor to become a middle layer of a three-tier π-stacking interaction between His-32 and Trp-33. Additional Van der Waals contacts from the Leu-94 side chain result in a highly complementary shape to the deepest part of the pocket. Two other important interactions are made with the backbone nitrogen atoms of Tyr-99 and Glu-100; the distance (2.73 and 2.77 Å) and angle between these atoms and the inhibitor carboxylate oxygens suggests the formation of strong hydrogen bonds (Fig. 5C).

Remaining contacts of inhibitor A with the pocket are less well defined hydrophobic interactions. The dichlorophenyl ring lies in a wider but shallower portion of the pocket without making obvious specific interactions. Although it has been found in previous studies (22) that both chlorine atoms contribute significantly to binding, neither atom is very close to the ε-nitrogen of His-32. In the complex structure, this interaction is lost and the rings of His-32 and Trp-33 become parallel to each other. Two other side chains, Leu-94 and Glu-100, also adjust their conformations to accommodate the inhibitor (Fig. 5C). Together, these changes in side chain conformation result in the formation of the observed binding pocket (Fig. 6).

TABLE III

| Scoring function | Inhibitor A | Inhibitor B |
|------------------|-------------|-------------|
| PMF              | −123        | −56         |
| PLP              | −88         | −24         |
| LUDI             | 1600        | 820         |

Three empirical scoring functions were applied to compound 1 bound in the location of either inhibitor A or B. For PMF (35) and PLP (36), a lower (more negative) score indicates better affinity, whereas for LUDI (37) a higher score predicts better affinity. For each program, scores are proportional to predicted binding energies.

In the apo structure, the aromatic rings of His-32 and Trp-33 form a T-shaped edge-to-face structure, with the hydroxyl group of Tyr-19 close enough to form a hydrogen bond with the e-nitrogen of His-32. In the complex structure, this interaction is lost and the rings of His-32 and Trp-33 become parallel to each other. Two other important interactions are made with the backbone nitrogen atoms of Tyr-99 and Glu-100; the distance (2.73 and 2.77 Å) and angle between these atoms and the inhibitor carboxylate oxygens suggests the formation of strong hydrogen bonds (Fig. 5C).

The ability of each protein to interact with E1 was scored qualitatively, because we did not control for variations in the level of protein produced in the in vitro transcription-translation reaction and, in any case, could not assess the proper folding or stability of the proteins. Most mutant proteins retained at least some ability to productively interact with E1 at the origin.

We also determined the sensitivity of these mutant proteins to inhibitor 2 (Fig. 3), an analogue of compound 1 that we have described previously (23). A number of substitutions had little if any effect on interaction with E1 or on inhibitor binding. These included V64A and K68A, which interact with inhibitor molecule B in the inhibitor-TAD complex, thereby providing further evidence that binding of this second molecule is not relevant to inhibition.

Substitutions I30V/M31L and I36M, which make the HPV11 TAD more similar to that of HPV6 E2, also had little effect on inhibition. These results suggest that none of these three residues are responsible for the 30-fold weaker affinity of HPV6 E2 for compound 2 (23). In further support that these residues do not play a major role in the type selectivity of the indandione inhibitors, we found that the corresponding mutations (V30I/L31M and M36I) in HPV6 E2 also had little effect on inhibition (data not shown).

Other substitutions (L15A, Y19A, E39A, and S98A) produced inactive or weakly active proteins. In principle this could indicate either that these side chains are directly involved in interaction with E1, or that the mutant proteins are not properly folded. The low signals obtained with the Y19A and E39A E2 proteins in the E1-E2-ori complex formation assay were insufficient to assess inhibition.

Substitution of Gln-12, His-32, Leu-43, Ser-65, and Ile-73 had a more deleterious effect on compound binding (IC_{50} val-

2 Davidson, W., McGibbon, G. A., White, P. W., Yoakim, C., Hopkins, J. L., Guse, I., Hambly, D. M., Frego, L., Ogilvie, W. W., Lavallée, P., and Archambault, J. (2004) Anal. Chem. 76, in press.
TABLE IV
Effect of amino acid substitutions in the E2 TAD on E1-E2-ori complex formation

| E2 mutation(s) | E1-E2-ori assembly | IC50
|---------------|---------------------|------|
| WT           | ++                  | ns   |
| Q12A         | ++                  | 20   |
| L15A         | +/−                 | 300  |
| Y19A         | −                   | 100  |
| I30/V/M31L   | ++                  | 20,000 |
| I36M         | +                   | 70   |
| E39A         | +/−                 | 40   |
| S40A         | +                   | 50   |
| L43A         | +                   | 130  |
| P60A         | +                   | 70   |
| V64A         | +                   | 70   |
| S65A         | +                   | 7,000 |
| K69A         | +                   | 140  |
| I73A         | +                   | 1,300 |
| L94A         | 200                 | 200  |
| S89A         | +/−                 | 240  |
| E100A        | ++                  | 8    |

TABLE V
Effect of selected amino acid substitutions on properties of E2 TAD proteins

| TAD mutations | Kd ( inhibitor 3) | IC50 E1 ATPase | I50 % E1 ATPase |
|---------------|-------------------|----------------|-----------------|
| WT            | 44.5 61 ± 8       | 4 ± 1          | 86 ± 4          |
| L15A          | 39.4 320 ± 30     | 6 ± 1          | 65 ± 1          |
| Y19A          | 44.6 120 ± 10     | 22 ± 3         | 32 ± 7          |
| E39A          | 12.0 100 ± 100    | 18 ± 2         | 38 ± 5          |
| I73A          | 46.2 410 ± 40     | 4 ± 1          | 90 ± 2          |
| E100A         | 42.4 5 ± 4        | 3 ± 1          | 83 ± 4          |

For E39A only a very weak signal was observed (S:B = 1.6) that was insufficient to evaluate inhibition by compound 2.

E2 TAD protein assembly reactions were carried out using in vitro translated HPV11 E2, E1-containing extract, and 3P-labeled origin DNA (see "Experimental Procedures"). A signal of 3100 cpm and signal: background (S:B) ratio = 31 was observed in assays performed with wild type E2. For mutant E2 proteins, E1-E2-ori complex formation was scored qualitatively as follows: ++ > 1000 cpm and S:B > 8; + = 500-1000 cpm and S:B = 3-8; +/− = 150-500 cpm and S:B = 2-3. IC50 values for compound 2 (Fig. 3) were determined by measuring complex assembly in the presence of compound concentrations ranging from 2 nM to 40 μM. Values shown are the average from duplicate determinations which in most cases differed by 30% or less.

Melting temperatures were determined by differential scanning calorimetry. Values shown are averages of two determinations that differed by less than 0.2 °C. Dissociation constants for compound 3 were determined using in vitro binding (see data in Fig. 7A). Parameters for inhibition of E1 ATPase activity were determined using a three-parameter logistic (WT, L15A, I73A, and E100A) or, when not possible, by estimation (Y19A and E39A), see data in Fig. 7B. For these experiments, values shown are the averages of triplicate or quadruplicate experiments ± SD.

Crystal Structure of the HPV E2-Inhibitor Complex

The remaining three mutant E2 proteins (L15A, Y19A, and E39A) differed from their WT counterpart in two ways. First, the maximal degree of E1-ATPase inhibition obtained at saturating concentrations of TAD was lower for these mutants than for WT E2. This was more prominent for Y19A and E39A E2,
Based on the structure of the HPV16 E2 TAD, Antson et al. (20) proposed the existence of an additional E2 dimerization interface within the TAD, comprised of the second and third α-helices, and notably involving a cluster of conserved residues (Arg-37, Ala-69, Ile-73, Gln-76, Leu-77, Glu-80, and Thr-81). Although the area of this interface is relatively large (~2000 Å²), it contains a number of water-mediated interactions and relatively few hydrophobic residue contacts, in contrast to most protein interfaces that are typically less hydrophilic than solvent-exposed surfaces (45, 46). Our observation that quite distinct monomer-monomer interactions are found in both of our HPV11 TAD structures, together with the results from chromatography and ultracentrifugation experiments that the HPV11 TAD is fully monomeric at concentrations up to ~20 μM calls into question the generality of the TAD dimerization interface and thus the likelihood that it plays a fundamental role in the function of E2, at least in the case of HPV11.

**Nature of the Inhibitor Binding Site**

The inhibitor binding pocket (Fig. 7) is not present in the E2 apo structure but rather appears to be induced upon inhibitor binding. This pocket is comprised of a large “shallow end” approximately defined by residues Gln-12, Glu-39, Lys-68, and Gln-71 and a smaller “deep end” defined by His-32, Trp-33, and Leu-94. The indandione moiety of inhibitor A is complementary to the deep end, and the resulting Van der Waals interactions, together with the hydrogen bonds between the carboxylate and backbone amides, account for the strongly exothermic binding of compounds in this series (Fig. 2C and Ref. 23). Additional interactions are made by the thiadiazoylphenyl moiety that extends out of the pocket, near Met-101.

**Conclusions Based on Studies with Mutant E2 Proteins**

**The Side Chain of Glu-100 May Impede Inhibitor Binding**

Mutation of Glu-100 to Ala did not affect substantially the interaction of E2 with E1, but specifically enhanced inhibitor binding as we reported previously (23). The structure of the inhibitor-TAD complex provides an explanation for this gain in affinity, namely that the entropic cost associated with the conformational change in the Glu-100 side chain necessary for compound binding (Fig. 5C) is likely reduced in the case of Ala-100.

**Amino Acids Outside the Pocket Can Reduce Inhibitor Binding**

A number of substitutions (Q12A, L15A, L43A, S65A, and I73A) clearly affected binding of the inhibitor, although they changed residues that do not directly contact the inhibitor in the structure. Further structural information would be required to fully understand the effects of these five substitutions, but all of these residues are located on helices in the three-helix bundle that defines the N-terminal subdomain of the TAD. The residues that form the deep pocket that binds tightly to the indandione moiety are also located within this three-helix bundle, and given the close fit between the inhibitor and this pocket it is possible that these substitutions subtly affect the packing of the three helices in a way that slightly alters the binding pocket. Perhaps similar subtle conformational changes explain why the indandione class of inhibitors binds to HPV6 E2 10–30-fold less tightly than to HPV11 E2, because the few amino acid differences in or near the binding pocket did not account for this type of differential activity. The effects of these non-binding site residue differences are conceptually similar to some HIV protease resistance mutations that occur well outside of the active site (47).
Importance of His-32 for Inhibitor Binding

We found that the specificity for inhibitor binding to low-risk rather than high-risk HPV types is in part because of a residue forming the pocket, His-32, which is a tyrosine in high risk types (Fig. 1). Substitution of His-32 for tyrosine in HPV11 E2 abrogated inhibitor binding, although it did not significantly affect interaction with E1. Based on our structure of the inhibitor-TAD complex, this substitution is predicted to prevent formation of the deep end of the inhibitor binding pocket. Furthermore, the converse experiment of inserting His-32 into HPV31 E2 generated a mutant protein that could interact with compound 2, albeit not to the same extent as HPV11 or HPV6 E2.

A Role for Tyrosine 19 in Binding to E1

Tyr-19 is highly conserved among HPV types although it is an Ile in bovine papillomaviruses. Substitution of this residue by Ala resulted in an HPV11 E2 protein that was still capable of binding the inhibitor and whose melting temperature and CD spectrum were very similar to that of wild type E2, indicating that this mutation does not affect the overall structure of the TAD. However, the Y19A mutant protein was defective in binding to E1 both in an E1-E2-ori complex formation assay and in an E1-ATPase inhibition assay. Thus, this mutation appears to specifically affect E1 binding. The proximity on the TAD surface of Tyr-19 to Glu-39, the only other residue specifically implicated in binding to E1, suggests that this surface of E2 might form part of the E1-E2 protein-protein interface.

Analysis of the two crystal structures presented here has provided a detailed understanding of the mechanism by which the indandione class of inhibitors antagonizes E1 binding by competing directly for the same or an overlapping binding surface on E2. In fact, the structure of an HPV18 E1-E2 TAD complex has recently been determined, and Tyr-19 does appear to make an important interaction with E1.4

Acknowledgments—We thank Dr. Michael Cordingley for stimulating discussions and Drs. Nathalie Goudreau, Christopher Lemke, and Steven Mason (Boehringer Ingelheim Canada Ltd.) as well as Prof. Liang Tong (Columbia University) for critical reading of the manuscript. We also thank Eric Abbate and Prof. Michael Botchan for communicating results prior to publication.

REFERENCES

1. Shah, K. V., and Howley, P. M. (1996) in Fields Virology (Fields, B. N., Knipe, D. M., and Howley, P. M., eds) pp. 2077–2109, Lippincott-Raven Publishers, Philadelphia, PA

2. Yao, J. M., Breiding, D. E., and Androphy, E. J. (1998) J. Virol. 72, 1013–1019

3. Benson, J. D., Lawande, R., and Howley, P. M. (1997) J. Virol. 71, 8041–8047

4. Rank, N. M., and Lambert, P. F. (1995) J. Virol. 69, 6323–6334

5. Li, R., Knight, J. D., Jackson, S. P., Tijan, R., and Botchan, M. R. (1991) Cell 65, 465–469

6. Spalholz, B. A., Vande Pol, S. B., and Howley, P. M. (1991) J. Virol. 65, 743–753

7. Breiding, D. E., Sverdrup, F., Grossel, M. J., Moscova, N., Boocock, W., and Androphy, E. J. (1997) Mol. Cell. Biol. 17, 7208–7219

8. Pursse, M., Ustav, E., Mandel, T., Stenlund, A., and Ustav, M. (1996) EMBO J. 15, 1–11

9. Lehman, C. W., and Botchan, M. R. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 4338–4343

10. Yang, L., Li, R., Mohr, I. J., Clark, R., and Botchan, M. R. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 6218–6225

11. Benson, J. D., and Howley, P. M. (1995) J. Virol. 69, 4364–4372

12. Berg, M. and Stenlund, A. (1997) J. Virol. 71, 3853–3863

13. LeMaster, D. M., and Richards, F. M. (1985) Biochemistry 24, 791–794

14. Benson, J. D., and Howley, P. M. (1995) J. Virol. 69, 4364–4372

15. Li, R., Knight, J. D., Jackson, S. P., Tijan, R., and Botchan, M. R. (1991) Cell 65, 465–469

16. Ferguson, M. K., and Botchan, M. R. (1996) J. Virol. 70, 6028–6032

17. Benson, J. D., Lawande, R., and Howley, P. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 4338–4343

18. Yang, L., Mohr, I. J., Botchan, M. R., and Pursse, M. (1996) EMBO J. 15, 2823–2829

19. Pursse, M., Ustav, E., Mandel, T., Stenlund, A., and Ustav, M. (1996) EMBO J. 15, 2823–2829

20. Antson, A. A., Burns, J. E., Moroz, O. V., Scott, D. J., Sanders, C. M., Bronstein, I. B., Dodson, G. G., Wilson, K. S., and Maitland, N. J. (2000) Nature 403, 865–869

21. Yang, L., Mohr, I. J., and Botchan, M. R. (1996) Science 278, 1673–1677

22. Giri, I., and Yaniv, M. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 6976–6980

23. Benson, J. D., Lawande, R., and Howley, P. M. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 4338–4343

24. Borgla, J. L., Philo, J. S., and MacBride, A. A. (1995) J. Virol. 70, 6028–6032

25. Pursse, M., Ustav, E., Mandel, T., Stenlund, A., and Ustav, M. (1996) EMBO J. 15, 2823–2829

26. Ferguson, M. K., and Botchan, M. R. (1996) J. Virol. 70, 6028–6032

27. Cooper, C. S., Umejeyi, S. N., and Winokur, P. L. (1998) Virology 241, 312–322

28. Wood, W. B. (1996) J. Mol. Biol. 16, 118–133

29. LeMaster, D. M., and Richards, F. M. (1985) Biochemistry 24, 7263–7268

30. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in Annual Report of the National Institute of Mental Health, National Institute of Mental Health, pp. 2077–2109, Lippincott-Raven Publishers, Philadelphia, PA

31. Shah, K. V., and Howley, P. M. (1996) in Fields Virology (Fields, B. N., Knipe, D. M., and Howley, P. M., eds) pp. 2077–2109, Lippincott-Raven Publishers, Philadelphia, PA

32. Benson, J. D., and Howley, P. M. (1995) J. Virol. 69, 4364–4372

33. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in Annual Report of the National Institute of Mental Health, National Institute of Mental Health, pp. 2077–2109, Lippincott-Raven Publishers, Philadelphia, PA
Crystal Structure of the E2 Transactivation Domain of Human Papillomavirus Type 11 Bound to a Protein Interaction Inhibitor

Yong Wang, René Coulombe, Dale R. Cameron, Louise Thauvette, Marie-Josée Massariol, Lynn M. Amon, Dominique Fink, Steve Titolo, Ewald Welchner, Christiane Yoakim, Jacques Archambault and Peter W. White

J. Biol. Chem. 2004, 279:6976-6985.
doi: 10.1074/jbc.M311376200 originally published online November 22, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M311376200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 47 references, 19 of which can be accessed free at http://www.jbc.org/content/279/8/6976.full.html#ref-list-1