Characterization of the Binding Site for Inhibitors of the HPV11 E1−E2 Protein Interaction on the E2 Transactivation Domain by Photoaffinity Labeling and Mass Spectrometry

Walter Davidson,[†,‡] Graham A. McGibbon,[*,†,§] Peter W. White,[§] Christiane Yoakim,[§] Jerry L. Hopkins,[‡] Ingrid Guse,[§] David M. Hambly,[§] Lee Frego,[‡] William W. Ogilvie,[§] Pierre Lavallée,[§] and Jacques Archambault[§]

Research and Development, Boehringer Ingelheim (Canada) Ltd., 2100 Cunard Street, Laval, Québec, Canada H7S 2G5, and Research and Development Center, Boehringer Ingelheim Pharmaceuticals Inc., 900 Ridgebury Road, Ridgefield, Connecticut 06877

An indandione-containing class of inhibitors abrogates DNA replication of human papillomavirus (HPV) types 6 and 11 by binding reversibly to the transactivation domain (TAD) of the viral E2 protein and inhibiting its interaction with the viral E1 helicase. To locate the binding site of this class of protein−protein interaction inhibitors, a benzophenone derivative was used to generate an irreversibly labeled E2-TAD polypeptide. The single site of covalent modification of the E2-TAD was identified by proteolytic digestions using trypsin, LysC, and V8 proteases and characterization of the resulting peptides by LC−MS procedures. Through this methodology, the benzophenone attachment point was located at the terminal methyl of residue Met101. Evidence further pinpointed the site of photoaffinity attachment to the terminal carbon atom, which is significant in providing a definitive example of the ability to locate photoinduced cross-linking to a polypeptide with atomic resolution using solely mass spectrometric detection. The location of the inhibitor binding site vis-à-vis the Glu39 and Glu100 residues sensitive to mutation for HPV 11 E2-TAD is discussed in relation to the crystal structure of the E2-TAD from the related HPV type 16.

Greater than 80 human papillomavirus (HPV) types have been identified that cause benign or malignant lesions of the epithelium, of which 25 infect the anogenital region. When not held in check by the immune system, their spectrum of associated disease ranges from genital warts to cervical lesions that can progress to invasive cancer. "Low-risk" types, HPV6 and HPV11, cause benign warts while cervical carcinomas and their precursor lesions are linked to infection by "high-risk" types, in particular HPV16, -18, -31, and -45.[1] The currently available therapeutic approaches for genital warts include a variety of ablative and cytodestructive treatments, including the use of liquid nitrogen, trichloroacetic acid, or topical administration of the immunomodulator imiquimod.[2] These treatments have drawbacks, and a need remains for an HPV-specific antiviral.

The circular DNA viral genome encodes six characterized early gene products and two capsid proteins. The only protein with enzymatic activity is the E1 helicase, which together with E2 is essential for viral DNA replication.[3] Most of the other early proteins are involved in regulating cellular proliferation and exert their effect by interacting with host cell factors.[4] Viral DNA replication is initiated by the binding of E1 and E2 to the HPV origin to form an E1−E2−origin complex. Assembly of this complex depends on the binding of E2, a sequence-specific DNA binding protein, to high-affinity sites located in the origin as well as on a crucial protein−protein interaction between E1 and E2. The latter serves to direct binding of E1 specifically to the origin since E1 binds DNA with little sequence specificity. The domains of E2 that bind to E1 and to the origin are located in the N-terminal and C-terminal portions of the protein, respectively, which are separated by a 70-amino acid "hinge" region of unknown function. The E1-binding domain of E2 is encoded within the first 200 amino acids, approximately, of the protein and is known as the transactivation domain (TAD) to reflect the fact that E2 also functions as a transcription factor to regulate expression of the viral genome. Assembly of the E1−E2−origin complex during the initiation of HPV DNA replication serves as a starting point for the assembly of additional E1 molecules into hexamers capable of unwinding the origin and the DNA ahead of the replication fork.[5] E1 and E2 also recruit to the origin the polymerase α primase and other host cell proteins required for replication.[6] As a consequence of their

* Corresponding author. Current address: Departments of Biochemistry and of Chemistry, McMaster University, Hamilton, ON, Canada. Fax: 905-522-2509. E-mail: mcgibbon@mcmaster.ca.

† Both authors contributed equally to this work.

‡ Boehringer Ingelheim Pharmaceuticals Inc.

§ Boehringer Ingelheim (Canada) Ltd.

(1) Shah, K. V.; Howley, P. M. In *Fields Virology*, 3rd ed.; Fields, B. N., Knipe, D. M., Howley, P. M., Eds.; Lippincott-Raven Publishers: Philadelphia, 1996; pp 2077−2109.

(2) Beutner, K. R.; Ferenczy, A. *Am. J. Med.* **1997**, *102*, 28−37.

(3) Chiang, C. M.; Ustav, M.; Stenlund, A.; Ho, T. F.; Broker, T. R.; Chow, L. T. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5799−5803.

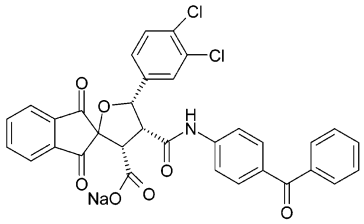
(4) Howley, P. M. In *Fields Virology*, 3rd ed.; Fields, B. N., Knipe, D. M., Howley, P. M., Eds.; Lippincott-Raven Publishers: Philadelphia, 1996; pp 2045−2076.

(5) Sanders, C. M.; Stenlund, A. *J. Biol. Chem.* **2000**, *275*, 3522−3534.

10.1021/ac035335o CCC: $27.50
Published on Web 00/00/0000

well-understood indispensability in replication, the proteins E1 and E2 are attractive targets for antiviral chemotherapy.[7,8]

Recently we described a series of compounds[9] that inhibit HPV DNA replication by interfering with the assembly of the E1−E2−origin complex. The first inhibitor in this class was identified by high-throughput screening of our corporate compound collection with an E1−E2−origin complex formation assay[10,11] and featured an indandione system spirofused onto an appropriately substituted tetrahydrofuran ring. It inhibited the protein−protein interaction between HPV11 E1 and E2 with an IC_{50} of 11 μM, by binding to the transactivation domain of E2 (E2-TAD),[12] the region of the protein that interacts with E1. Focused structural explorations[9] led to the discovery of more potent analogues that exhibited submicromolar IC_{50} values in the binding assay and activity in a cell-based assay of HPV DNA replication.[12] The information from structure−activity relationships, which indicated the viability of substitution at the amide attached to the tetrahydrofuran,[9,13] was employed to prepare **I**, a potent compound (IC_{50} of 610 nM)



Compound I

incorporating a photoreactive benzophenone appendage that could be exploited for photoaffinity labeling studies with the E2-TAD.

Benzophenones are well suited to photoaffinity labeling because they do not generate long-lived carbenes as reactive intermediates, which could lead to labeling at nonspecific binding sites.[14−17] Upon exposure to ultraviolet light, an excited-state radical intermediate is formed. To selectively label the protein target, an abstractible hydrogen atom is required in proximity to the benzophenone carbonyl oxygen atom (within 2.5−3 Å of the bound inhibitor).[18] After hydrogen transfer, the compound can cross-link to the protein via radical coupling, forming a carbon-based covalent bond that is generally stable through subsequent analyses.

Mass spectrometry has proven useful as an adjunct to radiometric or UV detection to aid in the identification of the site of photoaffinity labeling of a known polypeptide by a given ligand.[15,19−24] Recently, the effectiveness of online μLC−MS, coupling capillary-scale liquid chromatography with a magnetic sector mass spectrometer, to identify the site of photoaffinity ligand attachment to a specific amino acid residue, was demonstrated in one of our laboratories.[25] This analysis was performed without the requirement for special radioactive synthesis or reliance on LC detection of a ligand UV chromophore. This approach has several advantages in addition to speed and convenience. The online approach as described herein permits the detection of very low yield photoaffinity labeling and thus provides a semiquantitative method to investigate alterations in the incubation and photolabeling conditions in order to enhance yield. It is not plagued by ubiquitous problems of "false leads" caused by coelution of unlabeled peptides and breakdown or side products from the ligand, which are radioactive or have UV properties similar to those of a labeled peptide.

In this paper, we describe the photoaffinity labeling of a recombinant HPV11 E2-TAD protein with the benzophenone probe molecule, **I**, and the identification of the labeling site. The analysis consisted of LC−MS procedures plus trypsin, LysC, and V8 proteolysis, which eventually pinpointed the site of benzophenone binding to a single atom of the 24-kDa polypeptide. Instrumental for this characterization was the capability of capillary liquid chromatography combined with Fourier transform ion cyclotron resonance mass spectrometry (μLC−FTMS) to provide simultaneously high mass accuracy and resolution with good sensitivity.[26,27] The enhanced method described in the current study highlights the advantages of using FTMS to locate the site of photoaffinity attachment, in this particular example achieving atomic resolution solely using mass spectrometry detection.

EXPERIMENTAL SECTION

Materials. The probe molecule was synthesized from 3,4-dichlorophenylindanedione and 4-chlorobenzophenone as described for other E1−E2 interaction inhibitors.[9,13] Recombinant E2-TAD protein was produced in *Escherichia coli* and purified as previously described.[12] Proteolysis employed ammonium bicarbonate, DTT, and iodoacetamide from Sigma (Oakville, ON, Canada). Lyophilized sequencing grade trypsin, *Staphlococcus aureus* V8 protease, and endoproteinase LysC were obtained from Promega (Madison, WI). Reagent grade acetonitrile was obtained

(6) Bonne-Andrea, C.; Santucci, S.; Clertant, P.; Tillier, F. *J. Virol.* **1995**, *69*, 2341−2350.
(7) Phelps, W. C.; Barnes, J. A.; Lobe, D. C. *Antiviral Chem. Chemother.* **1998**, *9*, 359−377.
(8) Bernard, H. U. *Antiviral Ther.* **2002**, *7*, 219−237.
(9) Yoakim, C.; Ogilvie, W. W.; Goudreau, N.; Haché, B.; Naud, J.; O'Meara, J. A. *Bioorg., Med. Chem. Lett.* **2003**, *13*, 2539−2541.
(10) White, P. W.; Pelletier, A.; Brault, K.; Titolo, S.; Wardrop, E.; Thauvette, L.; Fazekas, M.; Cordingley, M. G.; Archambault, J. *J. Biol. Chem.* **2001**, *276*, 22426−22438.
(11) Titolo, S.; Pelletier, A.; Sauvé, F.; Brault, K.; Wardrop, E.; White, P. W.; Amin, A.; Cordingley, M. G.; Archambault, J. *J. Virol.* **1999**, *73*, 5282−5293.
(12) White, P. W.; Titolo, S.; Brault, K.; Thauvette, L.; Pelletier, A.; Welchner, E.; Bourgon, L.; Doyon, L.; Ogilvie, W. W.; Yoakim, C.; Cordingley, M.; Archambault, J. *J. Biol. Chem.* **2003**, *278*, 26765−26772.
(13) Yoakim, C.; Goudreau, N.; McGibbon, G. A.; O'Meara, J. A.; White, P.; Ogilvie, W. W. *Helv. Chim. Acta* **2003**, *86*, 3427−3444.
(14) Hatanaka, Y.; Sadakane, Y. *Curr. Top. Med. Chem.* **2002**, *2*, 272−288.
(15) Dormán, G.; Prestwich, G. D. *Trends Biotechnol.* **2000**, *18*, 64−77.
(16) Fleming, S. A. *Tetrahedron* **1995**, *51* 12479−12520.
(17) Brunner, J. *Annu. Rev. Biochem.* **1993**, *62*, 483−514.
(18) Dormán, G.; Prestwich, G. D. *Biochemistry* **1994**, *33*, 5661−5673.
(19) Brinegar, C. A.; Cooper, G.; Stevens, A.; Hauer, C. R.; Shabinowitz, J.; Hunt, D. F.; Fox, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5927−5931.
(20) Li, H.; Macdonald, D. M.; Hronowski, X.; Costello, C. E.; Leeman, S. E.; Boyd, N. D. *J. Biol. Chem.* **2001**, *276*, 10589−10593.
(21) Borchers, C. H.; Hochleitner, E. O.; Tomer, K.; Neamati, N. *Proceedings 50th ASMS Conference on Mass Spectrometry and Allied Topics*, 2002.
(22) Borchers, C. H.; Boer, R.; Klemm, K.; Figala, V.; Witte, S.; Denzinger, T.; Ulrich, W. R.; Hass, S.; Ise, W.; Gekeler, V.; Przybylski, M. *Mol. Pharmacol.* **2002**, *61*, 1366−1376.
(23) Jensen, O. N.; Steen, H. *Mass Spectrom. Rev.* **2002**, *21*, 163−182.
(24) Sachon, E.; Bolbach, G.; Chassaing, G.; Lavielle, S.; Sagan, S. *J. Biol. Chem.* **2002**, *277*, 50409−50414.
(25) Last-Barney, K.; Davidson, W.; Cardozo, M.; Frye, L. L.; Grygon, C. A.; Hopkins, J. L.; Jeanfavre, D. D.; Pav, S.; Qian, C.; Stevenson, J. M.; Tong, L.; Zindell, R.; Kelly, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 5643−5650.
(26) Emmett, M. R.; White, F. M.; Hendrickson, C. L.; Shi, D.-H. S.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 333−340.
(27) Martin, S. E.; Shabinowitz, J.; Hunt, D. F.; Marto, J. A. *Anal. Chem.* **2000**, *72*, 4266−4274.

Table 1. Sequences and Masses Pertaining to the HPV11 E2-TAD Polypeptide

residues	sequence			MW	A_0
tag+2−201[a]	GHHHHHH			23884.2	
	EAIAKRLDA	CQDQLLELYE	ENSIDIHKHI		
	MHWKCIRLES	VLLHKAKQMG	LSHIGLQVVP		
	PLTVSETKGH	NAIEMQMHLE	SLAKTQYGVE		
	PWTLQDTSYE	**M**WLTPPKRCF	KKQGNTVEVK		
	FDGCEDNVME	YVVWTHIYLQ	DNDSWVKVTS		
	SVDAKGIYYT	CGQFKTYYVN	FNKEAQKYGS		
	TNHWEVCYGS	TVICSPASVS	S		
121−147	FDG(CAM-C)EDNVMEYVVWTHIYLQDNDSWVK				3361.4696
85−108	TQYGVEPWTLQDTSYEMWLTPPKR				2925.4007
85−107	TQYGVEPWTLQDTSYEMWLTPPK				2769.2996

[a] The sequence of the E2-TAD (2−201) is depicted aligned but indicating the N-terminal His-tag (GHHHHHH) that is present to aid purification, but not in full-length E2 (http://us.expasy.org/cgi-bin/niceprot.pl?P04015). Residues not detected from the tryptic digest by LC−MS or μLC−FTMS are underlined and methionine 101 is indicated in boldface type.

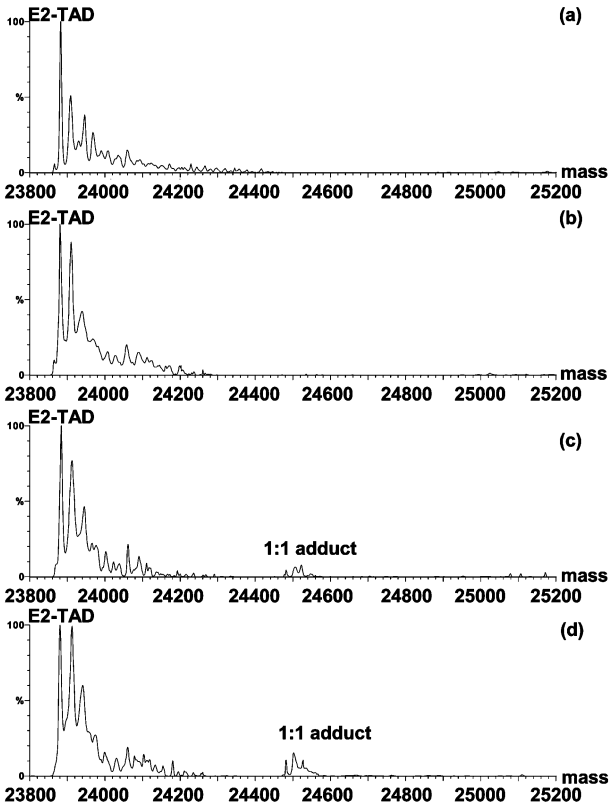


Figure 1. Deconvoluted mass spectra from samples of the E2 transactivation domain (E2-TAD) irradiated for the time indicated in the presence or absence of the photoreactive inhibitor: (a) 1 h without inhibitor, (b) 1 h with **I** and also a nonphotoreactive competitive inhibitor, (c) 1 h with **I**, and (d) 2 h with **I**. The masses measured for the E2-TAD ranged from 23 880 to 23 883 Da compared to the expected value of 23 884 Da. The adducts had observed masses of 24 502 and 24 516 Da as compared to 24 498 and 24 522 Da for E2-TAD + **I** in free acid and hydrated forms, respectively.

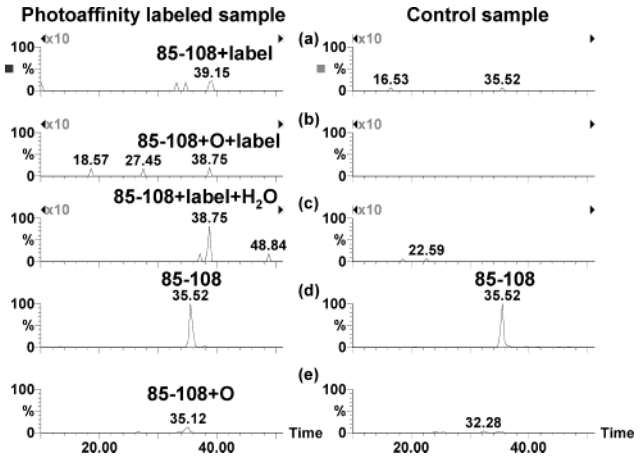


Figure 2. Comparison of μLC−FTMS ion chromatograms for ionized peptides from a sample of trypsin-digested E2-TAD that had been irradiated with inhibitor **I** to a control sample irradiated without it. Magnified 10× chromatograms corresponding to triply charged ions of residues (a) 85−108 + label (*m*/*z* 1180.498), (b) 85−108 + O + label (*m*/*z* 1185.830), (c) 85−108 + label + H_2O (*m*/*z* 1186.501), (d) 85−108 (*m*/*z* 976.141), and (e) 85−108 + O (*m*/*z* 981.473).

from EM Science (Gibbstown, NJ) and trifluoroacetic acid (TFA) was obtained from J.T. Baker (Phillipsburg, NJ).

Photoaffinity Labeling Experiments. The HPV11 E2-TAD polypeptide in buffer (20 μM in 20 mM Tris, pH 8.0, 100 mM NaCl, 0.1 mM EDTA) was incubated in the dark with a 1:1 molar ratio of inhibitor (or DMSO control) and then photolyzed for a total of 0−4 h at 4 °C using a Spectroline ENF-280C (Spectronics Corp.) lamp rated at 500 μW/cm^2 light intensity at a 6-in. distance from the sample. The samples were either exchanged into a solution of 80:20:1 $H_2O/CH_3CN/CH_3CO_2H$ or stored at −80 °C prior to further treatment. Initial proteolysis was conducted by adding a mixture of 8 M urea and 0.4 M ammonium bicarbonate in equal volume to the protein solution, reduction with DTT at 50 °C, alkylation with iodoacetamide at 50 °C, dilution, and then digestion to completion with enzyme (50:1 w/w, for 18 h at 37 °C).[28] The digests were acidified with TFA and, when not analyzed directly, stored at −80 °C. Further proteolysis of the trypsin digest solution with *S. aureus* V8 was conducted following initial LC−MS analysis. For this double digest, a 50-μL aliquot was adjusted back to pH 8 with 5 μL of 2.4 M ammonium bicarbonate and incubated 4 h after adding 0.25 μg of enzyme in 2.5 μL of the Promega resuspension buffer. An endoproteinase Lys C + V8 double digest of the labeled protein in 200 mM ammonium bicarbonate was performed, after reduction and alkylation, by simultaneously adding both enzymes to the sample. The sample was not acidified, and the portion not analyzed was immediately frozen.

(28) Stone, K. L.; Williams, K. R. In *A Practical Guide to Protein and Peptide Purification and Microsequencing*, 2nd ed.; Matsudaria, P., Ed.; Academic Press: San Diego, 1993; pp 45−69.

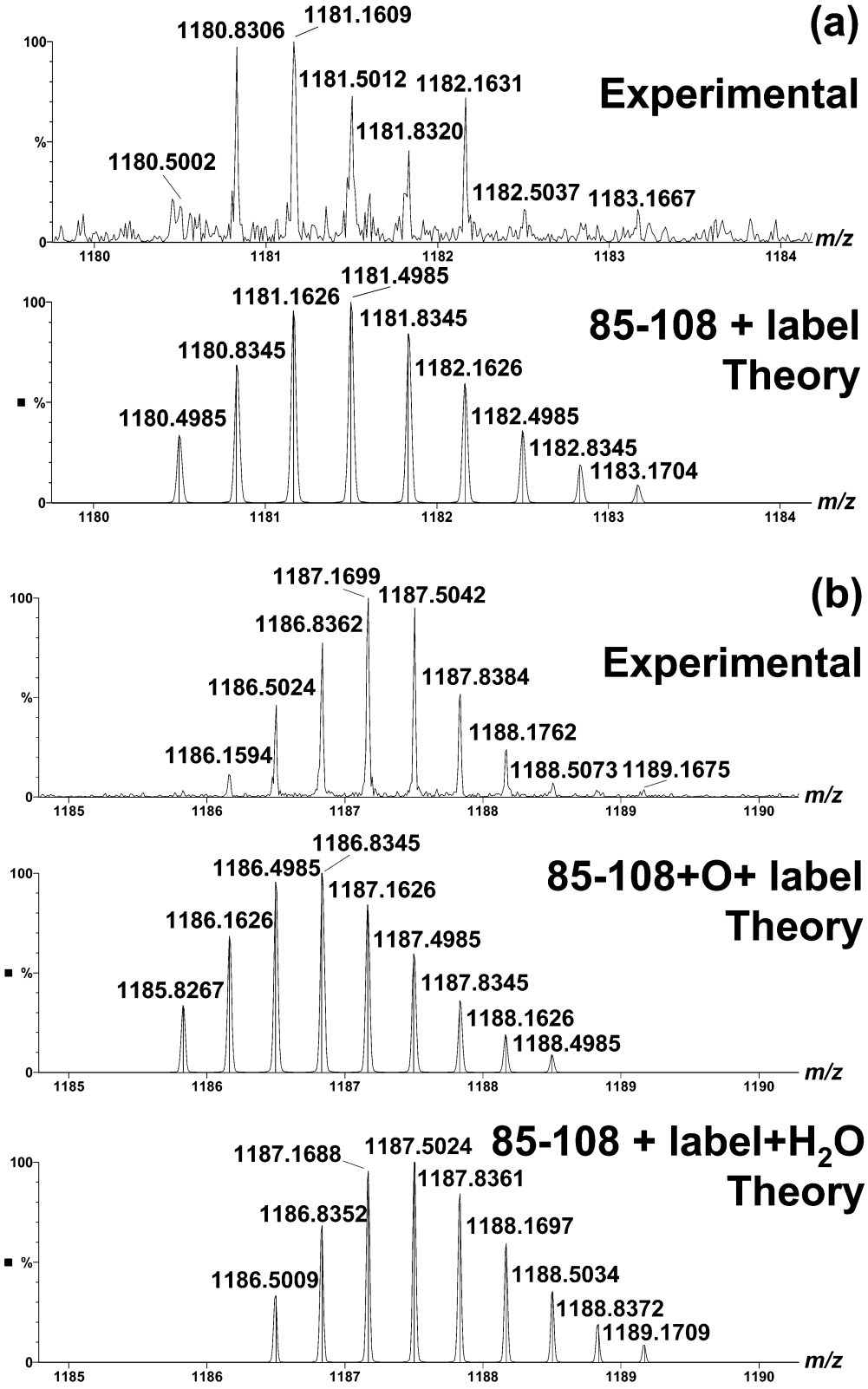


Figure 3. μLC−FTMS mass spectrum from photolabeled peptide ions obtained after trypsin proteolysis of E2-TAD that was irradiated with inhibitor **I**. Comparison of data is made to MassLynx-generated isotope distributions overlaid with centroid peaks for an ion corresponding to E2-TAD triply charged residues (a) 85−108 (TQYGVEPWTLQDTSYEMWLTPPKR) + label ($C_{33}H_{21}Cl_2NO_7$) and (b) the same labeled peptide either modified via oxidation, +O, or with a hydrated inhibitor, + H_2O.

Mass Spectrometry. LC−MS analysis of the protein samples was performed using a Symmetry C18, 2.1 × 100 mm, 5-μm column, a 2690 HPLC and 996 PDA (Waters), and a Quattro II mass spectrometer (Micromass) with electrospray ionization. Gradient chromatography employed CH_3CN/0.06% TFA and H_2O/0.06% TFA programmed from 20 to 60% organic in 30 min at

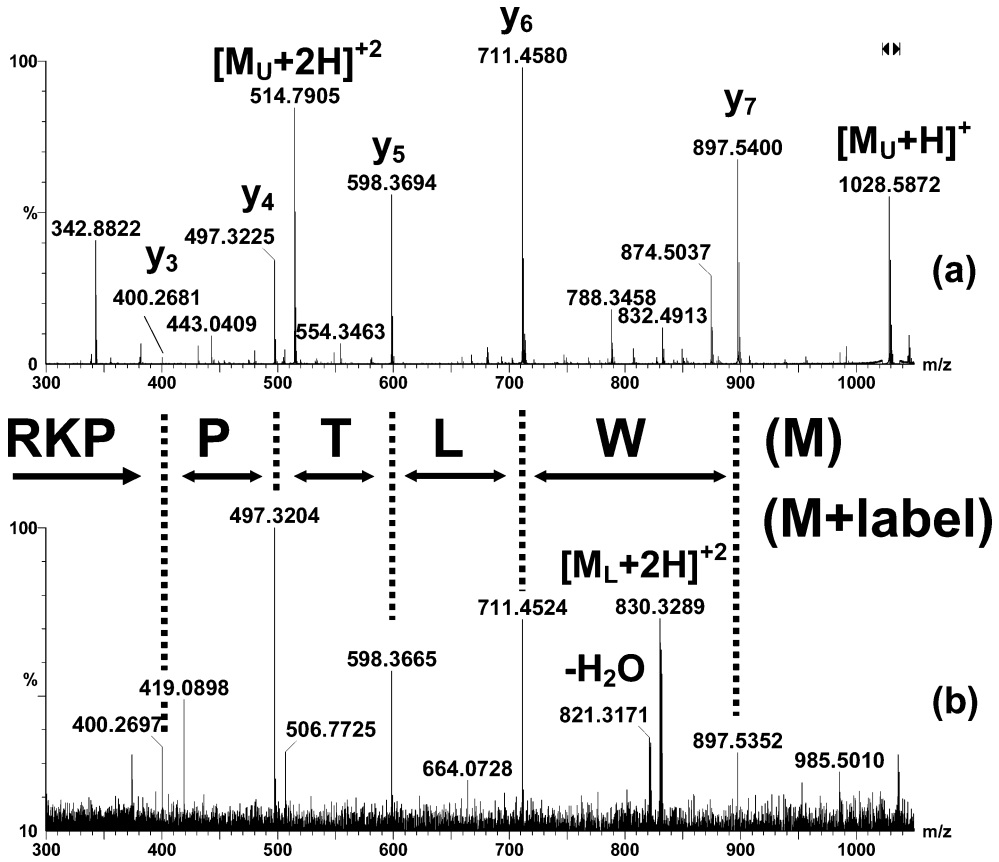


Figure 4. Collision-induced dissociation mass spectra of ions obtained from μLC−FTMS of trypsin/V8-digested E2-TAD, which had been irradiated with inhibitor **I**, that correspond to (a) M_U, the unlabeled peptide of sequence MWLTPPKR, and (b) M_L, the corresponding labeled peptide.

200 μL/min. Protein mass spectra were deconvoluted using Maximum Entropy software (Micromass).

Identification of the labeled tryptic peptide was performed on an Apex II 7-T Fourier transform mass spectrometer (Bruker). During chromatographic elution, data corresponding to 128 digitized spectra were acquired and employed to obtain survey data for detection of photoaffinity labeled peptides. Each spectrum was the average of four spectra of 2-s accumulation in the external hexapole. Data were acquired at 35 000 resolution. Capillary liquid chromatography was conducted with a 0.30 × 150 mm Pepmap C18 column (LC Packings, San Francisco, CA). A 20-μL injection volume was separated at 3 μL/min using mobile phase A, 99:1:0.1 $H_2O/CH_3CN/HCO_2H$, and mobile phase B, 10:90:0.1 $H_2O/CH_3CN/HCO_2H$, programmed from 0 to 100% B in a complex gradient over 75 min. The FTMS collision-induced dissociation experiments were conducted via in-source fragmentation at 190 versus 70 V for standard operation.

Xmass (Bruker) data files were converted to MassLynx (Micromass)-compatible format using custom software (Sierra Analytics) for data analysis. The data were analyzed exhaustively. The mass of the photolabel was added to each predicted tryptic peptide including peptides formed via missed cleavages. This was used to generate a list of m/z values for the charge states of each potential labeled peptide. The HPLC trace of the photolabeled tryptic digest was plotted for each theoretical m/z value. Positive results were compared to the control digest. Responses not present in the control were further analyzed by examining isotopic envelopes vis-à-vis peak spacing and abundances and checking for coelution of other charge states. Theoretical isotope distributions and exact masses were calculated for photolabeled peptides using MassLynx software. Accurate masses were determined via internal recalibration against coeluting peptides using the Xmass software and data sets.

RESULTS AND DISCUSSION

Demonstration of Protein Labeling. To produce a labeled polypeptide, the benzophenone inhibitor, **I**, was incubated with the HPV11 E2-TAD protein and irradiated for 1 h or more as described in the Experimental Section. Samples were analyzed by reversed-phase LC/UV/MS with electrospray ionization and found to contain two major components. The first-eluting peak produced a multiply charged spectrum deconvoluted to yield a mass of 23 880 ± 4 Da. This peak corresponds to the E2-TAD protein, which based on sequence has a predicted molecular mass of 23 884 Da (Table 1). The second component was ∼20% of the abundance of the first one by UV, and maximum entropy processing of the corresponding mass spectrum yielded a mass of 24 498 ± 10 Da, consistent with addition of **I** to the E2-TAD. The maximum entropy deconvoluted electrospray mass spectra of desalted sample mixtures (Figure 1) showed that adduct formation was specific, in that it was precluded by competition with a nonphotoreactive inhibitor and displayed dependence upon irradiation time, increasing in the 0−4-h time frame though further gains in adduct yield were not observed with extended photolysis. The spectra actually show a pair of peaks, consistent with the formation of E2-TAD labeled in part by **I** and also with the addition

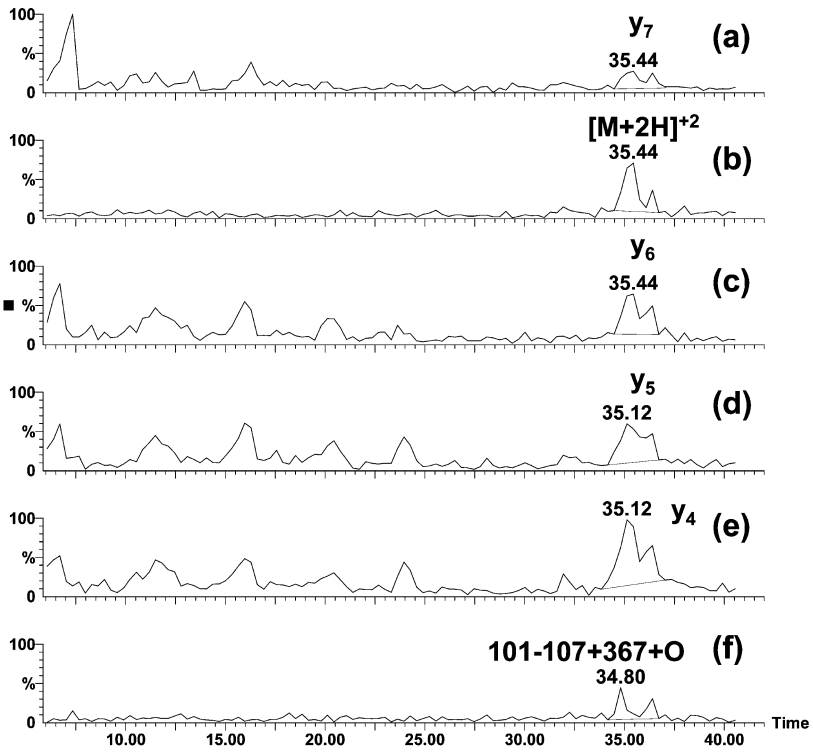


Figure 5. μLC−FTMS ion chromatograms for ionized peptides from a combination Lys C/V8 digest of E2-TAD that had been irradiated with inhibitor **I**. Chromatograms show coeluting ions corresponding to product ions from the modified peptide of residues 101−107 (MWLTPPK) + hydrated inhibitor (label + H_2O): (a) y_7, *m*/*z* 897.531; (b) $[M + 2H]^{2+}$, 830.330; (c) y_6, *m*/*z* 711.452; (d) y_5, *m*/*z* 598.368; and (e) y_4, *m*/*z* 497.320; as well as (f) ions corresponding to the oxidized peptide (MWLTPPK + O) + label, *m*/*z* 829.322.

of water because of the observed mass difference (24 504 vs 24 522 Da) and the known propensity of the indanediones to hydrate.[13]

Tryptic Digestion and Identity of the Labeled Fragment. Conditions were established wherein trypsin digestion was able to yield almost complete sequence coverage of unlabeled E2-TAD (Table 1). The partially labeled protein mixture was likewise denatured, reduced, and then alkylated with iodoacetamide after which it was enzymatically digested with trypsin. This mixture was analyzed by μLC/FTMS. To facilitate an exhaustive data analysis, the data were converted from the Bruker format to the MassLynx format. The LC/MS data were searched by generating exact mass chromatograms of all probable charge states for every predicted tryptic peptide, including those with one or two missed cleavage sites. The data were similarly searched for each peptide with the added mass of **I** ($C_{33}H_{21}Cl_2NO_7$). A weak response was generated for the triply charged peptide of residues 85−108 + **I** in the photoaffinity labeled sample (Figure 2a). This peak was present in the labeled sample but absent in a similarly prepared and analyzed control sample lacking **I** (Figure 2). Analysis of the full mass spectra at this retention time suggested additional overlapping components were present. Consistent with this, another small peak was identified in a chromatogram corresponding to the triply charged peptide of residues 85−108 + O + **I** (Figure 2b). A more abundant, but still weak (note the magnification 10× for labeled traces), response was obtained from a chromatogram of a triply charged ion consistent with a molecule cosisting of residues 85−108 + **I** + H_2O (Figure 2c). Both of these were also absent from the control sample (Figure 2). The unlabeled peptide was evident in both samples (Figure 2d) whereas its oxidation was apparently enhanced in the labeled sample (Figure 2e). Upon further investigation, signals corresponding to the doubly charged peptides were also found. These signals were present in the labeled sample but absent in a similarly prepared control sample lacking **I**. A close inspection of the mass spectrum from the labeled sample (Figure 3) showed *m*/*z* values and isotopic distributions consistent with the presence of **I** and its two Cl atoms. When the full spectra were recalibrated, using coeluting unrelated peptides as an internal mass standard, the monoisotopic masses obtained were consistent with the formula $C_{167}H_{217}N_{33}O_{47}SCl_2$ (and + H_2O/O) within 1 ppm. The ion intensity obtained for the labeled peptides was only a few percent of the corresponding unlabeled peptides. This low ion intensity could be attributed to reduced ionization efficiency or losses of the hydrophobic peptides in handling as is often encountered in LC analysis of photoaffinity labeled peptides.[14] Enrichment was considered because of the low yields and divided labeling encountered. However, due to the direct detection of adducts at high mass accuracy and resolution made possible with μLC−FTMS, there were no "false leads", i.e., other peptides, that actually had to be resolved with enrichment of the sample or competitive data prior to additional analyses. Note that V8 digestion alone of unlabeled E2-TAD also gave high sequence coverage but missed this critical stretch of residues.

Additional Digestion and Analysis. The low response and high mass of the labeled tryptic peptides made further resolution of the binding site problematic. The response was split between cleavage and missed cleavage products, oxidized or not, and further among a wide isotope distribution. High mass makes collision-induced dissociation more difficult and divides the signal among a larger number of fragment ions. For these reasons, an

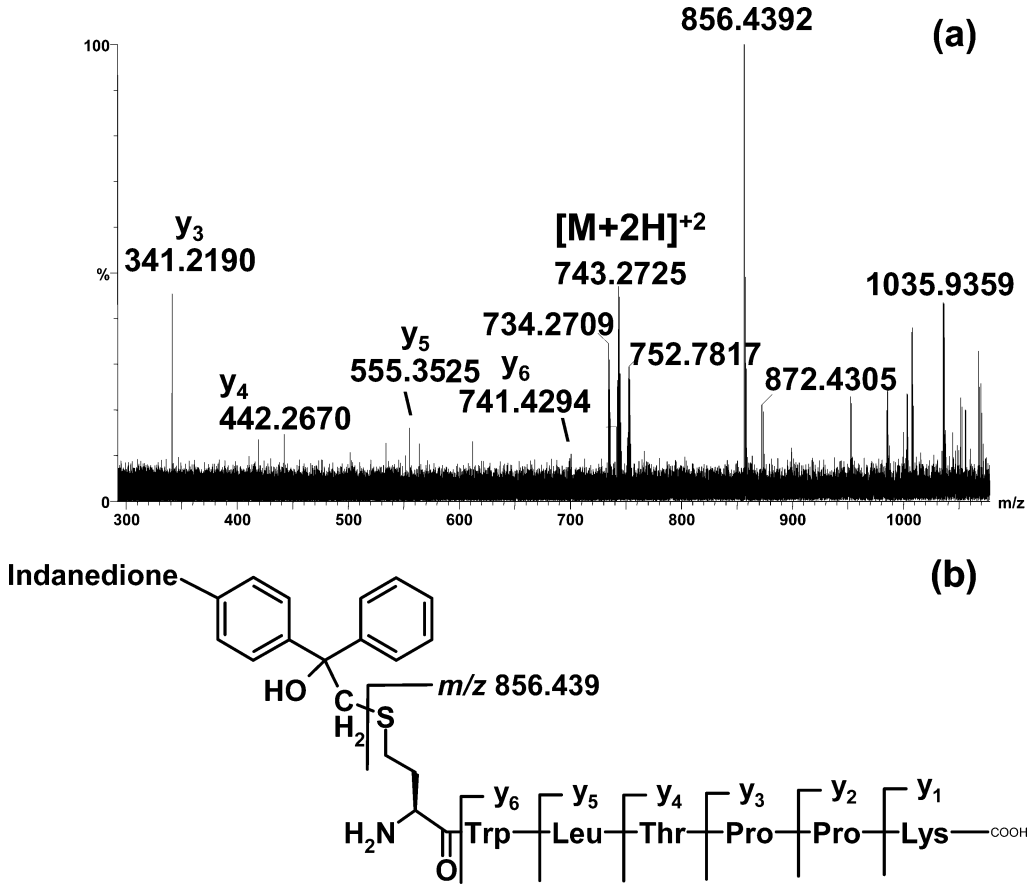


Figure 6. Labeled peptide from combined LysC/V8 proteolysis of E2 transactivation domain (E2-TAD): (a) μLC−FTMS collision-induced dissociation mass spectrum of labeled (and oxidized) peptide ions and (b) structure depicting proposed attachment of probe at the terminal methyl group of methionine 101. Peaks corresponding to the structure characteristic cleavages of the doubly charged ions are evident in the spectrum, particularly at the methionine (or methionine sulfoxide) sulfur atom with the difference between the accurately measured *m*/*z* 856.4392 and 872.4305 corresponding exactly to an atom of oxygen.

additional aliquot of labeled protein was subjected to digestion as described above, but with the addition of *S. aureus* V8 to the trypsin digest as a second enzymatic digestion step. A third aliquot was digested with two enzymes simultaneously, in this case endoproteinase LysC and V8, to avoid the partial missed cleavage occurring with trypsin as a consequence of the adjacent lysine and arginine residues (Lys107/Arg108).

Analysis of μLC/FTMS Data from Trypsin/V8 Digest. Analysis of the trypsin/V8 digest showed labeling of peptides 101−107 and 101−108. The original 85−107 and 85−108 peptides were no longer present, and no labeled peptide corresponding to residues 85−100 was observed. The μLC−FTMS instrument was operated with a higher "cone voltage" in the source in order to produce fragmentation of the peptides. This provides data to further resolve the site of photoaffinity attachment. Figure 4 shows the spectrum obtained for collisional fragmentation of 101−108. C-terminal fragment ions (y-series y_3, y_4, y_5, y_6, y_7) *without label attachment* are observed at low ppm deviation from their expected *m*/*z* values. $[M + H]^+$ and $[M + 2H]^{2+}$ ions with isotopic distributions and exact mass consistent with the modified 101−108 peptide (MWLTPPKR + **I** + H_2O) were observed. Coelution of other unrelated enzymatic fragments was observed. In particular, residues 121−147 in the +3 and +2 charge states were employed to provide an internal mass calibration. No b-series ions, labeled or unlabeled, were present. This was to be expected due to the location of positive charge at the C-terminus. No ions representing the peptide after loss of the label as a fragment or molecule, that is, *m*/*z* 1027 or *m*/*z* 1028, were observed, thus reducing concerns that the y-series ions might be formed via a double fragmentation of label located elsewhere in the sequence. The same y-series ions were produced from a peak eluting slightly later in the chromatogram, which was explicable as the peptide labeled without hydration. The labeled 101−107 peptide was also observed in this chromatogram with a weaker response. Analysis of the mass spectrum of labeled peptide produced by trypsin/V8 digestion thus verified the identification of the labeled peptide and located the site of label attachment to Met101.

Analysis of μLC−FTMS Data from LysC + V8 Digest. On examination of the μLC−FTMS ion chromatograms from the LysC + V8 double digest, a relatively intense response for labeled 101−107 (MWLTPPK + **I** + H_2O) was indeed observed (Figure 5). In this case (and in the labeling of 101−108 discussed above), under conditions conducive to fragmentation, the C-terminal fragment ions (y_3, y_4, y_5, y_6) all show a maximum at the same elution time as the singly and doubly charged molecules (Figure 5). This confirmed the location of label on the N-terminal Met101 as discussed above. The mass spectrum (Figure 6a) showed peaks for the unlabeled C-terminal fragment ions (y_3, y_4, y_5, y_6) and for loss of H_2O from the molecular ions. Of particular interest was the observation of an ion at *m*/*z* 856.4392. This was only consistent

with the original peptide sequence minus a methyl group (Loss of NH or O with addition of H would give a different exact mass.). The ion at m/z 872.4305 had the same nominal mass as unlabeled peptide $[M + H]^+$ but was inconsistent with its accurate mass. However, this fragment ion could be assigned as the loss of methyl from the same peptide sequence in methionine sulfoxide oxidized form, that is, via addition of O to the sulfur. This assignment was consistent with the observation of oxidation in the original tryptic peptides, i.e., doubly charged m/z 1471.71 ions. These observations permit the conclusion that the label was added to the side chain methyl of Met101, thus promoting cleavage of its bond to sulfur; the cleavage of this bond is enhanced in the small portion of the sample with oxidized methionine. The site of label attachment is thus defined as the side-chain methyl group of Met101 (Figure 6b).

Structural Insights. X-ray crystal structure data were reported for a portion of the HPV18 E2-TAD (residues 47−196) and the entire HPV16 E2-TAD (residues 1−201).[30] The existing structures provide sufficient framework in combination with the results reported here to identify a region of the HPV11 E2 protein capable of binding small molecules, which antagonize interaction with E1. As can be seen from Figure 7, the "L-shaped" E2-TAD is composed of two subdomains: N1, formed from three antiparallel α-helices, and N2, which is essentially an antiparallel β-sheet possessing a unique fold.[30] The E2-TAD of HPV11 shares 47% amino acid identity with HPV16 and 41% with HPV18, which are themselves 46% identical. At position 101, HPV16 has a valine and HPV18 a leucine, whereas several other HPV types including "low-risk" HPV6 and HPV11 have a methionine at this position. In this context, it should be emphasized that residue 101 represents the attachment point for the benzophenone probe rather than the binding site for the indanedione. Based on homology with HPV16, methionine 101 of HPV11 E2-TAD would lie at or near the junction of the two subdomains. Since this defines the benzophenone location, the rest of the inhibitor is limited to a position within ~200 pm of that spot, thereby placing the inhibitor binding site on the inside of the "L", opposite from the conserved residues in N1 that are implicated by mutagenesis as involved in transactivation (Arg37 and Ile73).[31] There is the possibility of extending the 195 pm toward Glu39 in N1, which is involved in replication as demonstrated by the fact its substitution by alanine abrogates

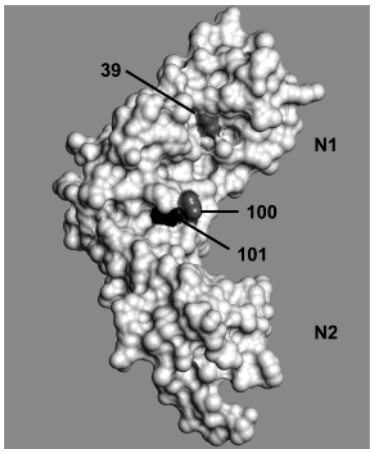


Figure 7. HPV16 E2-TAD surface modeled from the available X-ray crystal structure data (PDB entry 1dto)[30] showing the relative position of residue 101 (black), which is a valine instead of the methionine that gets labeled in HPV11. For orientation, the upper subdomain is N1 and the lower one is N2. The conserved Glu39 implicated in E1 interaction and Glu100 known to enhance sensitivity to the inhibitor class are indicated in dark gray. The surface was generated using the Insight II software package (Accelerys).

the E1−E2 interaction.[31] However, there was no pocket immediately visible on the HPV16 E2-TAD surface (Figure 7) capable of accommodating the indanedione inhibitors, which meant it was not obvious to pursue molecular modeling. Nevertheless, it was consistent with the fact that the indandiones are not active against HPV16 E2,[12] probably due to the limited sequence identity between both types. Prior investigation of the HPV11 E2-TAD revealed that mutation of Glu100 to Ala increased sensitivity to the indanedione inhibitors.[12] The location of residue 100 adjacent to the site of inhibitor cross-linking provides more direct evidence that this region of the E2-TAD is involved in inhibitor binding.

Conclusions. The concurrent high mass accuracy, resolution, and sensitivity attained using μLC−FTMS are particular advantages of the method described here, which enabled identification of photoaffinity labeled peptides that might otherwise have been missed due to low yield and distribution among modified ($+O/H_2O$) structures. The site of benzophenone attachment to the HPV11 E2 transactivation domain was determined by this method to be the terminal carbon atom of methionine 101, and the binding site for the indanedione class of inhibitors is therefore concluded to lie in this vicinity.

Received for review November 11, 2003. Accepted January 13, 2004.

AC035335O

(29) Harris, S. F.; Botchan, M. *Science* **1999**, *284*, 1673−1677.
(30) Anston, A. A.; Burns, J. E.; Moroz., O. V.; Scott, D. J.; Sanders, C. M.; Bronstein, I. B.; Dodson, G. G., Wilson, K. S.; Maitland, N. J. *Nature* **2000**, *403*, 805−809.
(31) Sakai, H.; Toshiharu, Y.; Benson, J. D.; Dowhanick, J. J.; Howley, P. M. *J. Virol.* **1996**, *70*, 1602−1611.