FAST COMMUNICATIONS

Acta Cryst. (1998). D**54**, 1471–1474

Expression, crystallization and preliminary X-ray analysis of the E2 transactivation domain from papillomavirus type 16

Julie E. Burns,[a] Olga V. Moroz,[b] Alfred A. Antson,[b] Cyril M. Sanders,[a] Keith S. Wilson[b]* and Norman J. Maitland[a] at [a]*Biology Department, York University, York Y01 5DD, England, and* [b]*Chemistry Department, York University, York YO1 5DD, England. E-mail: keith@yorvic.york.ac.uk*

(*Received 31 July 1998; accepted 13 August 1998*)

Abstract

The N-terminal transactivation domain of the E2 protein from human papillomavirus type 16 has been crystallized by vapour diffusion. Crystals belong to the space group $P3_121$ (or $P3_221$) with unit-cell dimensions $a = b = 54.3$, $c = 155.5$ Å. There is one molecule per asymmetric unit with a solvent content of 55%. Crystals diffract to at least 2.5 Å resolution and complete X-ray data to 3.4 Å have been collected on a conventional laboratory source. This 201-amino-acid domain of the E2 protein has been shown to interact functionally with both the HPV E1 protein and at least three cellular transcription factors, to fulfil its role in the control of viral transcription and replication. A knowledge of the structural basis of these multiple interactions should lead to a fuller understanding of the mechanism of action of this key regulator of the HPV life cycle.

1. Introduction

While the vast majority of human papillomaviruses (HPV) cause proliferative epithelial lesions such as warts, certain types are associated with the development of tumours, for example HPV-16, which is strongly implicated in the development of cervical cancer. Indeed, over 90% of cervical carcinomas contain HPV DNA from 'high risk' (for cancer) viruses (zur Hausen, 1996). In cancer cells, the viral genomes are normally integrated into the cellular chromatin at random human chromosomal locations, but the HPV E2 open reading frame is very frequently interrupted by this integration event (Choo *et al.*, 1987).

The E2 proteins of papillomaviruses are the major virally encoded regulators of HPV transcription and replication. E2 proteins are highly conserved nucleophosphoproteins with molecular weights of 45–48 kDa, which bind as dimers to specific DNA sequences found in the viral genome. The E2 proteins have two functional domains: an N-terminal domain of about 200 amino acids, which is highly conserved and functions in HPV gene transactivation and viral replication, and a C-terminal domain of about 70 highly conserved amino acids, responsible for DNA binding and dimerization. The domains are connected by a hinge or flexible linker of variable sequence (McBride *et al.*, 1991).

There are four E2 binding sites located in the upstream regulatory region (URR), of the HPV genome. This region also contains binding sites for a number of cellular transcription factors, and the p97 promoter, from which expression of the viral oncogenes E6 and E7 is initiated (O'Connor *et al.*, 1995). Expression of E6 and E7 is strongly implicated in the transformation of cells to malignancy (Munger *et al.*, 1989). However, experiments *in vitro* cannot confirm whether E2 acts as a repressor or activator of these genes (Phelps & Howley, 1987; Thierry & Howley, 1991; Bouvard *et al.*, 1994; Ushikai *et al.*, 1994; Dong *et al.*, 1994), but in cervical tissue samples E2 expression declines as early premalignant lesions progress to cancer (Stoler *et al.*, 1992; Maitland *et al.*, 1998).

Most studies of E2 function have used the E2 protein of bovine papillomavirus (BPV-1) as a prototype for all papillomaviruses, but important functional and possibly structural differences do exist between the human and bovine forms (Sanders *et al.*, 1995). The crystal structure of the BPV-1 E2 C-terminal DNA binding and dimerization domain has been previously determined (Hegde *et al.*, 1992),but to date no structure for the N-terminal domain has been reported. The gene transactivation and replication functions associated with this domain both require interaction with other viral (Storey *et al.*, 1995; Winokur & McBride, 1996; Yasugi *et al.*, 1997) or cellular proteins (Li *et al.*, 1991; Breiding *et al.*, 1997; Yao *et al.*, 1998). Determining the protein structure of this part of E2 may give important clues as to how it fulfils these functions.

2. Methods and results

2.1. *Cloning of E2 transactivation domain (E2NT)*

Codons 1–201 of HPV-16 E2 were amplified from CaSki cervical carcinoma cell line DNA by PCR with the following primers:

upstream 5' *CAAGACGTGCGCTAGCAT*-ATGGAGACTCTTTGCCA 3' (35mer)

NdeI

downstream 5' *CAGCAAGTGGATCCGCTAGCTTA*-GCTAAACACAGATGTAGGA 3' (42mer)

BamHI

(The bases in italics are synthetic linker and not homologous to HPV DNA.)

PCR products were purified using a Qiaquick PCR purification kit (Qiagen) and ligated directly into cloning vector pMOSBlue (Amersham). Inserts in pMOSBlue were sequenced and confirmed to be wild-type relative to the HPV-16 reference clone (Seedorf *et al.*, 1985; Meissner, 1997) before subcloning into the expression vector pET15*b* (Novagen). Clones containing inserts were sequenced to confirm that inserts were in frame with the N-terminal fusion 6x histidine-tag and transformed into *E. coli* expression strain BL21 (DE3) pLysS.

For protein expression, cultures derived from a single colony were grown in LB broth containing carbenicillin (50 µg ml^{-1})

© 1998 International Union of Crystallography
Printed in Great Britain – all rights reserved

ISSN 0907-4449 © 1998

and chloramphenicol (35 μg ml^{-1}) to an optical density of 0.5–0.6 (at 600 nm), induced with 1 m*M* IPTG and incubated at 298 K for 2.5 h. Cells were pelleted, washed in ice-cold 50 m*M* Tris–HCl, 50 m*M* NaCl, 2 m*M* EDTA, pH 7.9 at 277 K, re-pelleted and stored at 193 K.

2.2. *Purification of E2NT*

Pellets of induced bacteria were lysed and sonicated in ice-cold lysis buffer containing 40 m*M* Tris–HCl, 50 m*M* NaCl, 10 m*M* imidazole, 10% glycerol, 10 m*M* E64, 10 m*M* Pepstatin A, pH 7.9 at 277 K. The extract was adjusted to 0.5 *M* NaCl and centrifuged at 20 000*g*, 1 h, 277 K and the supernatant filtered through a 0.45 μm filter. E2NT was first purified through a 3 ml HisBind (Novagen) column, previously charged with 50 m*M* $NiSO_4$ and equilibrated in IMAC buffer (20 m*M* Tris–HCl, 0.5 m*M* NaCl, 10% glycerol, 5 m*M* imidazole, pH 7.9 at 277 K). Unbound protein was washed through with 10 column volumes (cv) equilibration buffer, while bound proteins were eluted with a step gradient using 4.5 cv each of IMAC containing 20, 40, 60, 100 and 120 m*M* imidazole. Remaining bound protein was eluted with 10 cv 20 m*M* Tris–HCl, 500 m*M* NaCl, 100 m*M* imidazole, 100 m*M* EDTA, 5% glycerol, pH 7.9 at 277 K. Samples from each fraction were analysed on 12% acrylamide–SDS gels containing 5 *M* urea. E2NT eluted in 100 and 120 m*M* imidazole fractions and elution buffer: these fractions were immediately adjusted to 10 m*M* DTT (dithiothreitol) and dialysed into 25 m*M* Tris–HCl, 40 m*M* NaCl, 5 m*M* DTT, 1 m*M* edta, pH 7.9 at 277 K.

Dialysed proteins were further purified through a 1 ml UNO Q (BioRad) column equilibrated in the same buffer and eluted on a 20 cv linear gradient of 40–600 m*M* NaCl in 25 m*M* Tris–HCl, 5 m*M* DTT, 1 m*M* EDTA, pH 7.9 at 294 K. Fractions containing E2NT, eluting at 320 m*M* NaCl, were pooled and concentrated using a Centriprep 10 concentrator (Amicon) before crystallization. Some preparations were further dialysed into 20 m*M* HEPES, 100 m*M* NaCl, 20% glycerol, 0.2 m*M* EDTA, 5 m*M* DTT, pH 7.9 at 277 K and stored frozen at 193 K. An example of the purification procedure is shown in Fig. 1. Final preparations were >98% homogeneous for the 23.5 kDa recombinant E2NT protein.

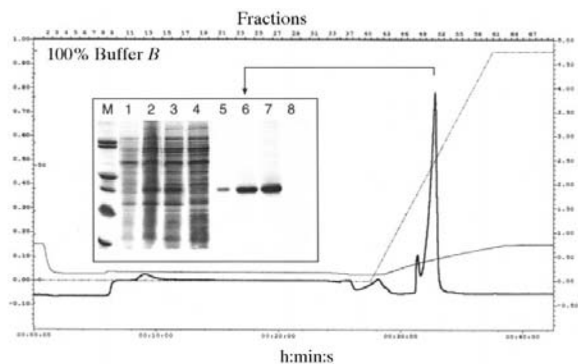

Fig. 1. Expression and purification of the bacterial E2NT domain showing the final Biologic (UNO-Q column) purification of E2NT. The *y* axis is uv absorbance (A_{280}) and *x* axis conductivity (mS cm^{-1}). The indicated peak is the desired product, which elutes over a narrow NaCl concentration range (5 × 1 ml fractions pooled at a salt concentration of 320 m*M*). The figure insert is an example of one such purification, as monitored by 12% SDS–urea–PAGE, stained with Coomassie brilliant blue. Lanes are as follows: lane *M*, BioRad low molecular weight range protein markers (97.4, 66, 45, 31, 21.5, 14.5 kDa, respectively, reading from the top of the gel); lane 1, uninduced BL21, DE3 +pET15*b* containing E2NT; lane 2, as 1 but induced by addition of IPTG; lane 3, lysate from 2 applied to primary His-Bind column; lane 4, unbound fraction from His-bind column; lane 5, pooled eluted bound fractions from His-bind column applied to UNO-Q column; lane 6, pooled peak fractions from UNO-Q column; lane 7, Centriprep 10 retentate: concentrated E2NT protein from UNO-Q column; lane 8, Centriprep 10 filtrate flow through.

2.3. *Crystallization*

Crystallization of the human papillomavirus E2 N-terminal transcriptional activation domain was carried out using the hanging-drop vapour-diffusion method. Prior to crystallization the purified protein was concentrated using a 10 K ultra-filtration membrane (Filtron) to 2.5–5.0 mg ml^{-1} in 10 m*M* Tris–HCl pH 8.0, 5 m*M* DTT, 0.2 m*M* EDTA, 300 m*M* NaCl. 2 μl drops of protein solution were mixed with 1 μl of the reservoir solution and equilibrated against 1 ml of the reservoir solution. The reservoir solution contained 0.8–1.2 *M* ammonium sulfate, 0.1 *M* triethanolamine pH 8.0–8.3 and 3–5% 2-methyl-2,4-pentanediol. Crystals (shown in Fig. 2) grew overnight at 292 K and reached a final size of 0.1 × 0.1 × 0.3 mm in 2–3 d. Crystals similar in size and shape could be obtained with 10–15% monomethylether polyethylene glycols (2000 or 5000), 0.1 *M* triethanolamine, pH range as above, 300 m*M* NaCl (NaCl was added to the wells after the drops had been mixed). Crystals did not grow at pH lower than 7.5, and were very small between pH 7.5 and 7.9.

Crystallization was only successful with very fresh protein preparations. After 3–4 d from final purification, protein samples no longer produced satisfactory crystals. This is probably reflected in the limited lifetime of the crystals themselves (described below).

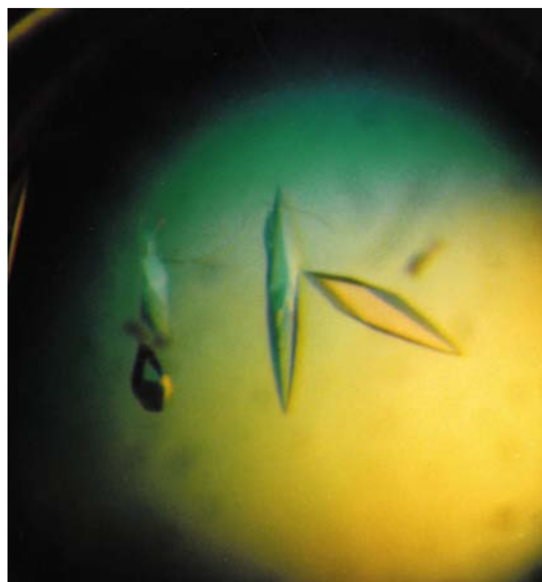

Fig. 2. Crystals of the N-terminal transactivation domain of the E2 protein from human papillomavirus type 16 grown with monomethylether polyethylene glycol 5000. Approximate size of the crystals is 0.1 × 0.1 × 0.3 mm.

Table 1. *Statistics of the X-ray data*

Resolution range (Å)	No. of unique reflections	$\langle I/\sigma(I)\rangle$	$I/\sigma(I) > 3$ (%)	Completeness (%)	$R(I)$† (%)
30.00–7.29	455	20.0	90.5	97.7	4.3
7.29–5.80	410	11.4	82.4	97.9	8.5
5.80–5.07	401	9.9	79.8	98.5	9.3
5.07–4.61	383	10.1	79.8	97.3	9.6
4.61–4.28	394	9.1	76.0	97.5	10.5
4.28–4.03	382	7.3	73.9	98.2	13.8
4.03–3.83	384	5.6	61.2	96.1	18.5
3.83–3.66	365	4.6	56.4	96.3	23.4
3.66–3.52	390	3.6	50.2	98.3	29.4
3.52–3.40	364	3.4	48.5	96.5	28.8
Overall	3928	9.1	70.4	97.4	10.9

† The merging $R(I)$ factor is defined as $I = \sum |I - \langle I\rangle| / \sum I$.

2.4. *X-ray data*

The crystals were characterized using Cu $K\alpha$ X-ray radiation from a Rigaku RU200 rotating-anode generator with MAR Research image-plate scanner as a detector. The crystals were frozen at 120 K using a cryocooling system (Oxford Cryosystems Cryostream). Crystals grown from ammonium sulfate were transferred into the cryoprotectant solution containing 40% ammonium sulfate, 20% glycerol, 0.1 *M* triethanolamine pH 8.3. The crystals grown from PEG 5000 monomethylether were soaked for 30 s in cryosolution containing 65% PEG 5000, 0.1 *M* triethanolamine pH 8.3, 300 m*M* NaCl.

Crystals belong to the space group $P3_121$ (or $P3_221$) with cell dimensions $a = b = 54.3$, $c = 155$ Å. Assuming one 23.5 kDa protein molecule in the asymmetric unit the specific volume V_M is 2.4 Å^3 Da^{-1} (Matthews, 1968) and the solvent content is 55%. The crystals diffract to at least 2.5 Å, Fig. 3, and complete data have been collected to 3.4 Å resolution on the above source. Data were processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). 11 015 individual measurements were reduced into 3928 unique reflections giving an average redundancy for each reflection of 2.8. The data are 97.4% complete with an overall merging *R* factor ($RI = \sum |I - \langle I\rangle| / \sum I$) of 10.9%. The statistics of the X-ray data set are summarized in Table 1. A search for the heavy-atom derivatives is in progress. Synchrotron radiation will be used to extend the resolution of the data.

As stated above, crystals could only be grown from very freshly prepared protein samples. In addition, after growing to maximum size in 2–3 d, the crystals deteriorated in terms of diffraction quality in less than one week. This problem was addressed by cryogenic freezing and storage of crystals immediately after growth. A comparable approach will also be necessary for the heavy-atom derivative search. E2 is a sensitive and inherently flexible protein. It proved very difficult over several years to create a construct which was amenable to crystallization. However, recent advances in cryogenic freezing have proved essential to this project, which would have been intractable only a few years ago.

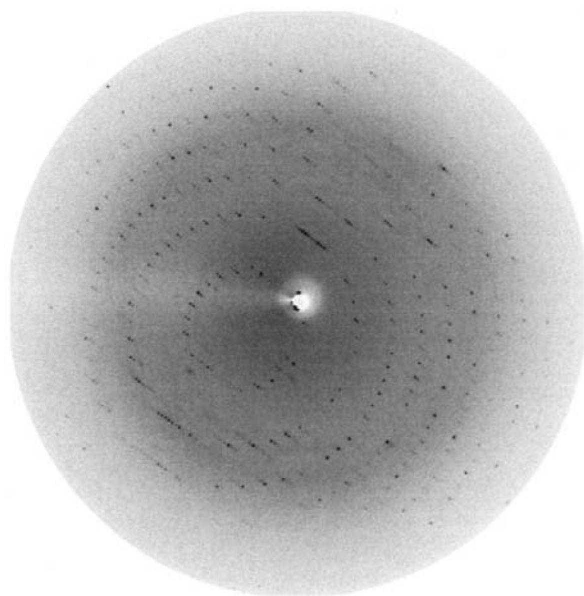

Fig. 3. A 0.1° oscillation image of an E2 transactivation domain crystal grown with monomethylether polyethylene glycol 5000. The image was recorded using a 300 mm Mar Research image plate. The resolution at the edge of the image is 2.5 Å.

We thank Yorkshire Cancer Research and the BBSRC for financial support. We also thank Lucy Hopwood, Paula Daniels for technical assistance and Mike Hodkinson for help in the X-ray data collection.

References

Choo, K. B., Pan, C. C. & Han, S. H. (1987). *Virology*, **161**, 259–261.

Bouvard, V., Storey, A., Pim, D. & Banks, L.(1994). *EMBO J.* **13**, 5451–5459.

Breiding, D. E., Sverdrup, F., Grossel, M. J., Moscufo, N., Boonchai, W. & Androphy, E. J. (1997). *Mol. Cell. Biol.* **17**, 7208–7219.

Dong, G., Broker, T. R. & Chow, L. T. (1994). *J. Virol.* **68**, 1115–1127.

zur Hausen, H. (1996). *Biochim. Biophys. Acta*, **1288**, F55–F78.

Hegde, R. S., Grossman, S. R., Laimins, L. A. & Sigler, P. B. (1992). *Nature (London)*, **359**, 505–512.

Li, R., Knight, J. D., Jackson, S. P., Tjian, R. & Botchan, M. R. (1991). *Cell*, **65**, 493–505.

McBride, A. A., Romanczuk, H. & Howley, P. (1991). *J. Biol. Chem.* **266**, 18411–18414.

Maitland, N. J., Conway, S., Wilkinson, N. S., Ramsdale, J., Morris, J. R., Sanders, C. M., Burns, J. E., Stern, P. L. & Wells, M. (1998). *J. Pathol.* In the press.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Meissner, J. (1997). In *Human Papillomaviruses 1997: A Compilation & Analysis of Nucleic Acid & Amino Acid Sequences*, edited by G. Myers, F. Svendrup, C. Baker, A. A. McBride, K. Munger, H.-U. Bernard & J. Meissner. New Mexico: Los Alamos National Laboratory of Theoretical Biology & Biophysics.

Munger, K., Phelps, W., Bibb, V., Howley, P. & Schlegel, R. (1989). *J. Virol.* **63**, 4417–4421.

O'Connor, M., Chan, S.-Y. & Bernard, H.-U. (1995). In *Human Papillomaviruses 1995: A Compilation of Nucleic Acid & Amino Acid Sequences*, edited by G. Myers, H.-U. Bernard, H. Delius, C. Baker, J. Icenogel, A. Halpern & C. Wheeler. Los Alamos, New Mexico: Theoretical Biology & Biophysics, Los Alamos National Laboratory.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Phelps, W. C. & Howley, P. M. (1987). *J. Virol.* **61**, 1630–1638.

Sanders, C. M., Stern, P. L. & Maitland, N. J. (1995). *Virology*, **211**, 418–433.

Seedorf, K., Krammer, G., Durst, M., Suhai, S. & Rowekamp, W. G. (1985). *Virology*, **145**, 181–185.

Stoler, M. H., Rhodes, C. R., Whitbeck, A., Wolinsky, S. M., Chow, L. T. & Broker, T. R. (1992). *Hum. Pathol.* **23**, 117–127.

Storey, A., Piccini, A., Massimi, P., Bouvard, V. & Banks, L. (1995). *J. Gen. Virol.* **76**, 819–826.

Thierry, F. & Howley, P. M. (1991). *New Biologist*, **1**, 90–100.

Ushikai, M., Lace, M. J., Yamakawa, Y., Kono, M., Anson, J., Ishiji, T., Parkkinen, S., Wicker, N., Valentine, M.-E., Davidson, I., Turek, L. P. & Haugen, T. H. (1994). *J. Virol.* **68**, 6655–6666.

Winokur, P. L. & McBride, A. A. (1996). *Virology*, **221**, 44–53.

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

Yasugi, T., Benson, J. D., Sakai, H., Vidal, M. & Howley, P. M. (1997). *J. Virol.* **71**, 891–899.