Protein and Peptide Letters, Vol. 8, No. 4, pp. 323-326, 2001
Bentham Science Publishers Ltd.
0929-8665/01 $ 20.00 + 0.00

CRYSTALLIZATION AND X-RAY DATA ANALYSIS OF THE EXTENDED DNA-BINDING DOMAIN OF THE E2 BOVINE *PAPILLOMAVIRUS* TYPE 1 PROTEIN

João A. R. G. Barbosa[a], Carlos Basílio Pinheiro[a], Adriana C. Alves[a], José R. Brandão Neto[a], Leonardo G. Alonso[b], Gonzalo de Prat-Gay[b] and Igor Polikarpov[a*]
[*] Corresponding author. E-mail: igor@lnls.br, fax:+55 (19)3287-4632

[a] Laboratório Nacional de Luz Síncrotron, Caixa Postal 6192, CEP 13084-971, Campinas - SP, Brazil
[b] Fundación Campomar, Av. Patricias Argentinas 435 (1405), Buenos Aires, Argentina

Abstract
Hanging-drop vapour-diffusion method was used for the crystallization of the extended DNA-binding domain (residues 309-410) of the E2 protein from Bovine *Papillomav*irus Type 1. X-ray data collection at 2.0 Å resolution was performed using synchrotron radiation. The crystal symmetry could be described by the space group $P3_121$ and with unit cell parameters a = b = 55.3 Å, c = 203.4 Å. The protein structure was solved by molecular replacement.

1. Introduction

Papillomaviruses are double stranded DNA viruses that cause epithelial lesions, often developing into malignant transformation, in particular cervical cancer [1]. The genome of the Bovine *Papillomavirus* Type 1 (BPV1) codes for eight different proteins. The E2 protein is the one responsible for the regulation of gene transcription and participates in viral genomic DNA replication [2]. Three domains can be distinguished in E2: activation domain, flexible "hinge" domain and DNA-binding domain. The E2 DNA-binding domain (E2-C) is located at the C-terminal end of the polypeptide, from residue 326 to 410. E2-C is active as a homodimer that binds to a conserved viral palindromic sequence. Each monomeric unit contains two α-helices and four β-strands. In the dimer the eight β-strands, four from each monomer, form a barrel. The most important regions of E2-C involved in the DNA binding are located in the first α-helix and in the loop between the second and third β-strands [3,4]. Recently the NMR structure of the extended E2-C domain with an extra 16-residue segment (residues 310 to 325) has been determined [5]. Although no homology was identified in this region among over

70 different *Papillomaviruses*, the extra 16 residues in the BPV1 E2-C domain have been implicated in enhancing the stability of the DNA binding domain [5].

In the present work we report the crystallization and the preliminary X-ray structure analysis of the extended BPV1 E2-C domain. The protein was produced from an engineered gene and contains 17 extra residues (residues 309 to 325) at the N-terminal when compared with the minimal DNA-binding domain comprising the E2-C residues 326 to 410. The 17 residues added to the N-terminal of the E2-C minimal DNA-binding domain corresponds to a portion of the flexible "hinge" between transactivation and DNA-binding domains. The protein was crystallized by hanging-drop vapour diffusion method. The space group $P3_121$ describes the crystal symmetry. The protein structure was determined by molecular replacement and the asymmetric unit of the crystal contains two functional dimers of the extended BPV1 E2-C domain.

2. Methods, results and discussion

Figure 1 - Extended BPV1 E2-C domain crystal.

Cloning, expression and purification of the E2-C domain will be described elsewhere (Leonardo G. Alonso and Gonzalo de Prat-Gay, unpublished results). The initial concentration of the protein solution used in this work was 8.64 mg/ml in 25 mM TRIS pH 7.5, 100 mM NaCl and 10 mM DTT. Crystallization was achieved by screening suitable conditions with the sparse-matrix screen (Crystal Screen I, Hampton Research Corp.) using hanging-drop vapour diffusion at 291 K. The drops had 2 µl and contained 1:1 parts of reservoir and protein solutions. Refinement of screen condition number nine, including variation in pH, temperature and in polyethylene glycol (PEG) concentration, yielded crystals with dimensions up to 0.2 x 0.2 x 0.1 mm^3 after one week (Figure 1). The best crystals were obtained with the reservoir solution contained 0.1 M sodium citrate pH 4.6 to 5.2, 28-32% PEG 4000 and 0.2 M ammonium acetate. Changing the temperature to 277 K did not improve crystallization.

Data collection was performed at the Protein Crystallography Beamline of the Brazilian National Synchrotron Light Laboratory (LNLS), Campinas, Brazil [6,7]. The best resolution achieved with the MAR345 imaging plate was 2.0 Å using 1.39 Å synchrotron radiation. Auto-indexing, integration, scaling and merging were performed using the HKL package [8]. All saturated reflections were included in the scaling and all the overlapping reflections were excluded from it. The crystal parameters and data collection statistics are listed on Table 1. An example of the diffraction pattern is shown in Figure 2.

Using Matthews's volume calculations [9], the asymmetric unit was expected to contain 4 monomers (V_m = 2.35 $Å^3Da^{-1}$), equivalent to two functional dimers. Solution of the structure was accomplished by molecular replacement technique using crystal structure of the E2-C dimer from 2BOP (PDB code) [3] as a search model and the CNS program version 1.0 [10]. Using data from 4 to 10 Å resolution the correct solution had a correlation coefficient of 0.56 and a R-factor equal to 0.39.

In a previous 2.5 Å resolution BPV1 E2-C structure refinement in the *R32* space group [4] three dimers were found in the asymmetric unit. Amino acid residues Cys40, putatively involved in the redox regulation of the DNA binding by BPV1 E2-C [11], were

Temperature (K)	100
Space group	*P3₁21*
Cell parameters (Å)	*a*=55.3; *c*=203.4
Resolution (Å)	30.0 - 2.0
Number of Images	145
Oscillation range (°)	1.0
Mosaicity range (°)	0.36-0.4
Image plate distance (mm)	160.0
Number of measurements	127541
Number of unique reflections	24850
Completeness (%)	99.8 (98.9)
Mean multiplicity	5.1 (5.3)
Mean I/σ(I)	16.7 (3.1)
R_{merge} (%)	8.7 (62.0)

Table 1- Data collection statistics. Numbers in parenthesis are for the highest resolution shell, 2.05 to 2.00 Å.

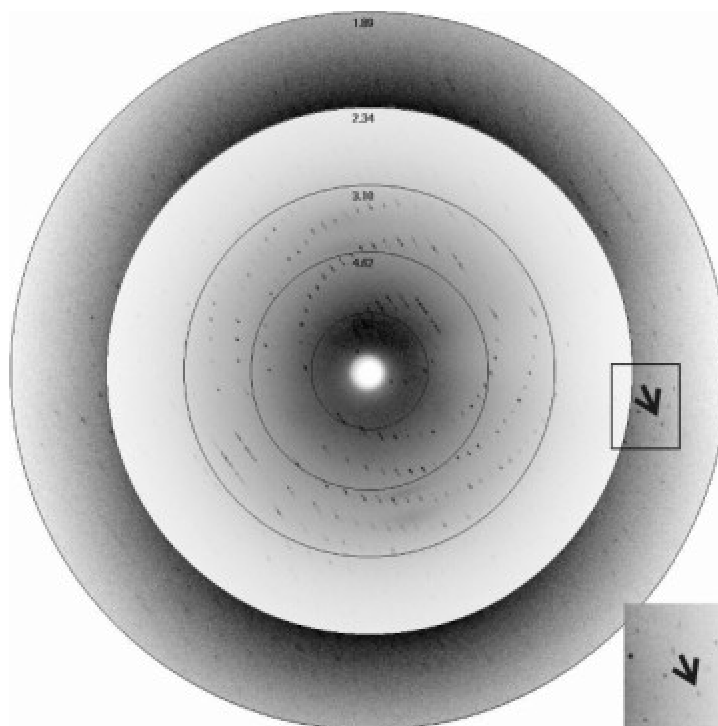

Figure 2 – The diffraction pattern of the extended BPV1 E2-C crystal. Numbers are the resolutions of the rings in Å. The outer region (> 2.34 Å) is 5.5 times more saturated than the inner regions (< 2.34 Å). Arrows are marking the same reflection at 2.0 Å resolution.

found to form disulphate bonds responsible for dimers linkage in the asymmetric unit cell. These disulphate bonds not exist in solution under physiological conditions and are, probably, a consequence of the crystallization under non-reduced conditions. Analysis of the molecular replacement solution clearly indicates that in the present crystal form the extended BPV1 E2-C domain does not make disulphate bridges between the functional dimers in the asymmetric unit. The closest distance between the thiol groups (Cys227-Cys356) is

approximately 10 Å. The distance between Cys40 from different dimers is about 15 Å. The model is now being further refined.

Acknowledgements We would like to thank Gustavo H. N. Moreira for computer administration. Financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, grant numbers 98/12043-4, 99/03387-4, 99/07150-9 and 99/08042-5), Fundación Antorchas and Convênio Argentino-Brasileño de Biotecnología is acknowledged.

References

1. Howley, P. M. (1996). Papillomaviridae: The viruses and their replication *in Virology*, 2045-2076, Lippincott-Raven, New York.
2. McBride, A. A, Romanczuk, H. & Howley, P. M. (1991). *J. Biol. Chem.* **266**, 18411-18414.
3. Hegde, R. S., Grossman, S. R., Laimins, L. A. & Sigler, P. B. (1992). *Nature* **359**, 505-512.
4. Hegde, R. S., Wang, A. F., Kim, S. S. & Schapira, M. (1998). *J. Mol. Biol.* **276**, 797-808.
5. Veeraraghavan, S., Mello, C. C., Androphy, E. J. & Baleja, J. D. (1999). *Biochemistry* **38**, 16115-16124.
6. Polikarpov, I., Oliva, G., Castellano, E. E., Garratt, R., Arruda, P., Leite, A. & Craievich, A. (1997). *Nucl. Instrum. Methods A* **405**, 159-164.
7. Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. & Craievich, A. (1997). *J. Synchrotron Rad.* **5**, 72-76.
8. Otwinowski, Z. (1993). Oscillation data reduction program *in Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs S. Bayley, pp. 56-62, SERC Daresbury Laboratory, UK.
9. Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491-497.
10. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905-921.
11. McBride, A. A., Klausner, R. D. & Howley, P. M. (1992). *Proc. Natl. Acad. Sci. USA*, **89**,7531-7535.