doi:10.1016/j.jmb.2006.03.056 J. Mol. Biol. (2006) 359, 667–678

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Structure of the Epstein-Barr Virus Oncogene BARF1

Nicolas Tarbouriech[1,2], Florence Ruggiero[3], Mireille de Turenne-Tessier[4] Tadamasa Ooka[4] and Wim P. Burmeister[1,2,5]*

[1]*EMBL Grenoble outstation BP181, F-38042, Grenoble cedex 9, France*

[2]*Institut de Virologie Moléculaire et Structurale FRE 2854, CNRS—UJF, BP181 F-38042, Grenoble cedex 9 France*

[3]*Institut de Biologie et Chimie des Protéines, UMR 5086 CNRS—UCB, 7 passage du Vercors, F-69367 Lyon Cedex 07 France*

[4]*Laboratoire de Virologie Moléculaire, UMR 5537 CNRS—Faculté de Médecine R.T.H. Laennec, Rue G. Paradin F-69372 Lyon cedex 08, France*

[5]*Institut Universitaire de France, 103, bd Saint-Michel F-75005 Paris, France*

**Corresponding author*

The Epstein-Barr virus is a human gamma-herpesvirus that persistently infects more than 90% of the human population. It is associated with numerous epithelial cancers, principally undifferentiated nasopharyngeal carcinoma and gastric carcinoma. The BARF1 gene is expressed in a high proportion of these cancers. An oncogenic, mitogenic and immortalizing activity of the BARF1 protein has been shown. We solved the structure of the secreted BARF1 glycoprotein expressed in a human cell line by X-ray crystallography at a resolution of 2.3 Å. The BARF1 protein consists of two immunoglobulin (Ig)-like domains. The N-terminal domain belongs to the subfamily of variable domains whereas the C-terminal one is related to a constant Ig-domain. BARF1 shows an unusual hexamerisation involving two principal contacts, one between the C-terminal domains and one between the N-terminal domains. The C-terminal contact with an uncommonly large contact surface extends the beta-sandwich of the Ig-domain through the second molecule. The N-terminal contact involves Ig-domains with an unusual relative orientation but with a more classical contact surface with a size in the range of dimer interactions of Ig-domains. The structure of BARF1 is most closely related to CD80 or B7-1, a co-stimulatory molecule present on antigen presenting cells, from which BARF1 must have been derived during evolution. Still, domain orientation and oligomerization differ between BARF1 and CD80. It had been shown that BARF1 binds to hCSF-1, the human colony-stimulating factor 1, but this interaction has to be principally different from the one between CSF-1 and CSF-1 receptor.

Keywords: Epstein-Barr virus; BARF1; immunoglobulin domain; CD80; B7-1

Introduction

The Epstein Barr virus (EBV) is a human gamma-herpesvirus that persistently infects more than 90% of the world population. If the first infection occurs early during childhood, it often goes unrecognized but EBV can lead to infectious mononucleosis if the primo-infection occurs during adolescence or adulthood. The virus persists life-long in B cells and is able to promote B cell immortalization *in vitro*. In immunocompromised and immunosuppressed patients it can lead to lymphoproliferative disease due to an absence of the control of the virus by the immune system. EBV is not only one of the causative agents of Burkitt's lymphoma, a childhood cancer common in parts of Africa where *Plasmodium falciparum* malaria is endemic,[1] but EBV is also associated with numerous human epithelial cancers including undifferentiated nasopharyngeal carcinoma (NPC) and gastric carcinoma. NPC is a major health problem in South-East Asia and North Africa, with an incidence of 5–40 per 100,000 individuals.[2–4] Its association with EBV is constant except in a few atypical highly differentiated cases.[5,6] About 5–20% of gastric carcinoma, a widespread cancer, are associated with EBV.[7]

Among the about 86 proteins encoded by the EBV genome,[8] two viral oncogenes, LMP1 and BARF1, were shown to induce malignant transformation in rodent fibroblasts.[9–11] LMP1 and BARF1 are respectively expressed in 50% and 90% cases of NPC.[12,13] BARF1 transcripts are detected in a high proportion of EBV-positive gastric carcinomas,

Abbreviations used: EBV, Epstein-Barr virus; Ig, immunoglobulin.

E-mail address of the corresponding author: wpb@embl-grenoble.fr

where LMP1 expression is consistently negative.[14] BARF1-encoded protein was also found in epithelial cells immortalised by NPC-derived EBV[15] and in NPC biopsies.[12] The oncogenic activity of the BARF1 gene was first demonstrated in rodent fibroblasts.[10] BARF1 protein can also immortalize primary monkey kidney epithelial cells[16] and induce malignant transformation in other established cell lines, such as human Louckes B cells[17] and EBV-negative AKATA[18] cells. BARF1-immortalised epithelial cells are not tumourigenic in nude mice, while BARF1-transfected immortalised cell lines induce tumours.[10,17] This suggests that BARF1 can intervene in two oncogenic processes: immortalisation and malignant transformation. Several cellular genes (Bcl2, CD21, CD23 and CD71) are transactivated in BARF1-transfected cells.[11,17,18] Following transfection of rodent fibroblast or simian epithelial cell lines, or after infection of human epithelial cells with recombinant adenovirus, or upon viral cycle activation in EBV-positive B cell lines the BARF1 protein is secreted into the medium after signal sequence (residues 1–20) cleavage.[19–21]The addition of purified BARF1 protein into serum-free culture medium of Balb/c3T3 fibroblasts, human B cells and primary monkey epithelial cells resulted in cell cycle activation,[20] suggesting that BARF1 protein can act as a cell growth factor.

Moreover, it has been shown also that BARF1 binds to the human cytokine hCSF-1 and reduces the action of hCSF-1 on the proliferation of macrophages[19] and the alpha interferon production by mononuclear cells,[22] suggesting that BARF1 protein acts also in immunomodulation.

The BARF1 gene, its protein and its function have recently been reviewed.[23] A BARF1 homologue with 75% sequence identity is found in Cercopithecine herpesvirus 15 (Rhesus lymphocryptovirus, rLCV,[24]), a lymphocryptovirus closely related to EBV present in old world primates. As LCV is strictly species-specific and as there is no BARF1 homologue present in LCV of new world primates it seems to be acquired late in evolution after the separation of old and new world primates.[25]

Here we report the crystal structure of the BARF1 protein purified from human cells at 2.3 Å resolution. The structure presents two immunoglobulin (Ig) domains in a unique hexameric assembly.

Results

Protein production, crystallization and structure determination

The BARF1 protein was expressed in HeLa cells using a recombinant adenovirus system. As it is secreted, it was purified from the culture media by concanavalin A affinity chromatography followed by size-exclusion chromatography where the protein eluted as an oligomer with an approximative molecular mass of 240 kDa. Crystals were first obtained by the vapour diffusion method using nanodrop crystallization trials using standard screens. Initial conditions were manually refined using the hanging-drop method in order to obtain crystals of suitable size for data collection. Rhombohedric crystals belonging to spacegroup $H3$ with cell parameters $a=b=179.25$ Å, $c=95.72$ Å diffracted to 2.3 Å. Despite systematic twinning of the crystals by merohedry with twinning fractions between 0.1 and 0.45 the structure was solved by the single anomalous dispersion (SAD) method using a platinum derivative of a crystal with a twinning fraction of 0.12. All steps of the structure determination used less than 0.5 mg of pure protein. The final model was refined to an R-factor of 17.5% ($R_{free}=23.4\%$). The asymmetric unit contains four protein chains with the visible residues 21 to 161 and 173 to 220 organized in two dimers. Residue 21 is the first residue of secreted BARF1 after cleavage of the signal peptide. Statistics of the data collection and refinement are shown in Table 1.

Table 1. Data collection and model refinement statistics

	BARF1 + Pt
Data collection statistics	
Wavelength (Å)	0.9393
Resolution (Å)	30–2.3 (2.42–2.3)
Unique reflections	50,910
Completeness (%)	100 (100)
R_{sym} (%)	5.9 (22.8)
$I/\sigma I$	10.7 (3.1)
Multiplicity	5.3 (4.7)
Refinement statistics	
R_{cryst} (%)	17.5 (17.9)
R_{free} (%)	23.4 (27.0)
Mean atomic B-factor (Å^2)	28.1
RMS on bond length (Å)	0.022
RMS on bond angles (deg.)	2.15
Model composition (in the asymmetric unit)	
Residues	748
Water molecules	533
Platinum atoms	12
N-acetyl glucosamine	(Nag)
Mannose (Man)	4

Values for the highest resolution bin are given in parentheses.

Overall fold

The BARF1 protein structure is composed of two domains belonging to the immunoglobulin fold superfamily, the first one ranging from residues 21 to 123 and the second one from residues 125 to 220 (Figure 1). The N-terminal domain is clearly related to the variable subfamily (V-set) with an ABED/C″C′CFG organization where the A strand is hydrogen bonding to the G strand rather than the B strand (Figure 1(b)). The C-terminal domain is more difficult to classify due to the disordered region 162–172, potentially containing the strand C′ or D. Nevertheless, an ABED/CFG organization characteristic of the constant subfamily (C-set) is most probable as the few residues visible in electron