Protein Expression and Purification 67 (2009) 35–40

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression of Epstein-Barr virus EBNA1 protein in *Escherichia coli*: Purification under nondenaturing conditions and use in DNA-binding studies

Naïma Bouallag, Claire Gaillard, Vincent Maréchal, François Strauss *

Centre de Recherche des Cordeliers, INSERM—Université Pierre et Marie Curie—Université Paris Descartes, Paris, France

ARTICLE INFO

Article history:
Received 14 January 2009
and in revised form 3 April 2009
Available online 22 April 2009

ABSTRACT

Epstein-Barr virus nuclear antigen 1 (EBNA1) is a viral protein required for stable replication and segregation of DNA episomes containing the Epstein-Barr virus (EBV) origin of replication, *OriP*. Overproduction of EBNA1 protein in *Escherichia coli* has previously been shown to be difficult due to the large number of codons in EBNA1 gene that are infrequently used in *E. coli*. Here we changed the 26 rare codons that are found among the first 78 codons of EBNA1 gene, and replaced them with codons that encode the same amino-acids but are abundant in *E. coli*. This led to a significant improvement of EBNA1 expression in a standard *E. coli* strain. Partial EBNA1 polypeptides of 11.5–16 kDa extending from the N-terminus to the second arginine and glycine-rich region were extremely abundant in the extract, however, resulting in a second limitation of the level of EBNA1 expression. EBNA1 was expressed as a fusion with a C-terminal six-histidine tag in order to get rid of the short polypeptides by Ni–NTA affinity purification, and salt conditions were used that allowed us to extract and purify EBNA1 without resorting to protein denaturing reagents. The purified protein was used in DNA-binding experiments with DNA fragments containing specific EBNA1 sites. The *E. coli*-expressed protein formed specific DNA–protein complexes that could be analyzed in polyacrylamide gels without showing the aggregation, or linking, phenomenon that is usually observed with EBNA1 expressed in eukaryotic cells. EBNA1 protein expressed in *E. coli* should therefore prove useful to further study the biochemical properties of this crucial Epstein-Barr virus protein.

Introduction

During latent infection, the replication of the genome of Epstein-Barr virus (EBV)[1] requires only two viral elements: the virus-encoded EBNA1 protein (Epstein-Barr virus Nuclear Antigen 1), and the origin of replication *OriP*, a ~2000 bp region of the viral DNA genome that contains several specific EBNA1-binding sites. All other elements of the replication system are provided by the host cell, and episomal DNA vectors containing the *OriP* sequence can be replicated and stably maintained through generations in human cells expressing EBNA1 [1]. Therefore, in addition to its importance in EBV function, the *OriP*–EBNA1 system has great potential in gene therapy [2,3].

Analyzing the characteristics of the *OriP*–EBNA1 system is therefore of great interest. The *OriP* sequence contains two distinct domains, the FR region that is formed of a series of 20 tandem repeats of the EBNA1-binding site regularly spaced by 30 bp, and the DS region that contains the actual replication origin, with four EBNA1-binding sites. EBNA1 is a dimeric protein composed of two copies of a polypeptide of 641 amino-acids (EBV strain B95-8), that contains several distinct domains which have been mapped in great detail (see e.g. [1] for a recent review of the *OriP*-EBNA1 system). Despite the large amount of data that has been obtained about this system, the exact mechanism of EBNA1-mediated replication and maintenance of *OriP* episomes is not fully understood. In particular, no enzymatic activity of EBNA1 has yet been found. As EBNA1 is only expressed at very low levels in human cells, most biochemical studies have relied on heterologous expression systems.

The C-terminal DNA-binding domain, which is responsible for the dimerization and the site-specific DNA-binding of EBNA1, can be very efficiently expressed in *Escherichia coli* and purified. Its biochemical properties and its interactions with DNA have been thoroughly studied, as well as its crystal structure either alone or as a cocrystal in association with its specific DNA binding site [4,5]. In contrast, expressing full length EBNA1 in *E. coli* has proven technically difficult, and for that reason EBNA1 has most often been studied using proteins expressed in insect cells with the baculovirus system and purified by a succession of chromatographic steps, a successful but expensive and time-consuming procedure. It will therefore be very useful to be able to express EBNA1 in bacteria. Recently, Duellman and Burgess [6] noted that the EBNA1 gene possesses a large number of codons that are infrequently used in *E. coli*. It is known that the presence of rare codons in a coding sequence can severely affect protein expression due to a lack of cog-

* Corresponding author.
E-mail address: francois.strauss@crc.jussieu.fr (F. Strauss).

[1] *Abbreviations used:* EBV, Epstein-Barr virus; EBNA1, Epstein-Barr virus nuclear antigen 1; IPTG, isopropylthiogalactoside; GuHCl, guanidine hydrochloride; EMSA, electrophoretic mobility shift assay; PCR, polymerase chain reaction; DRiPs, defective ribosomal products.

1046-5928/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2009.04.009

nate tRNAs that would be required for an efficient translation. This is particularly noticeable when rare codons are present in blocks, or when they are clustered near the translation start site of mRNAs [7,8]. For example the arginine codons AGG and AGA, which are very common in EBNA1 coding sequence, are read by $tRNA_4^{Arg}$ that is present only at very low levels in *E. coli*. Duellman and Burgess [6] were thus able to express EBNA1 by using an *E. coli* strain harboring a plasmid that expresses genes for rare tRNAs. This system allowed the same authors to express and purify very large amounts of EBNA1 [9]. By comparing the expression levels obtained with EBNA1 with levels obtained with other proteins, however, one would consider that it should be possible to further increase EBNA1 expression in *E. coli*. In addition, understanding the mechanisms that limit EBNA1 expression in *E. coli* may eventually bring information about the biochemical function of EBNA1 itself. Here we studied whether EBNA1 expression could be further improved by replacing the rare codons that code for 26 of the first 78 amino-acids of the protein with codons that encode the same amino-acids but are abundant in *E. coli*. At the same time we observed a very high abundance of partial EBNA1 polypeptide chains, extending from the N-terminus to a region that was mapped by mass spectrometry and found to lie within the second glycine and arginine-rich region at positions 330–375 of the EBNA1 amino-acid sequence. After expressing EBNA1 fused to a C-terminal tag, we worked out conditions for its extraction and its purification without using protein-denaturing chemicals. We used the purified protein in electro-mobility shift assay (EMSA) experiments with DNA fragments containing different numbers of EBNA1-binding sites, including a fragment from the DS region that contains four specific EBNA1 sites. We found that the protein expressed in *E. coli* forms specific complexes with DNA without showing the aggregation or linking phenomenon that has been observed with EBNA1 expressed in other cells.

Materials and methods

Plasmids

EBNA1 protein contains a large domain, composed exclusively of glycine and alanine residues, that can be deleted without affecting the function of the protein in the replication and segregation of *OriP* episomes. Therefore all constructs used in the present work, which contain a deletion of most of the glycine–alanine domain, will be noted EBNA1. Cloning vectors were pET15b and pET22b (Novagen), which are designed for cloning genes downstream of a T7 polymerase promoter for overexpression in *E. coli* upon induction of T7 RNA polymerase with IPTG. Proteins expressed from pET15b vectors have a six-histidine tag at their N-terminus, proteins expressed from pET22b vectors having a six-histidine tag at their C-terminus.

Plasmids used for EBNA1 expression in *E. coli* were constructed by classical cloning techniques, and were verified by sequencing. The constructs used are represented in Fig. 1. Each vertical bar on the maps of the first two constructs, A and B, represents a codon of low frequency usage in *E. coli*.

Construct A, pET15b-EBNA1, contains the wild-type EBNA1 nucleotide sequence (with amino-acids 90–322 deleted in the glycine–alanine region) cloned in expression vector pET15b.

Construct B, pET15b-$EBNA1_{(MC15-78)}$, expresses exactly the same protein as the previous construct, but all the 26 rare codons found from position 15–78 in the EBNA1 nucleotide sequence were replaced by codons that are common in *E. coli*. At the same time, all codons with frequencies lower than average were also replaced with common codons. Forty-three out of 64 codons were thus changed in this region using appropriate synthetic oligonucleotides. The modified nucleotide sequence has been deposited in GenBank under Accession No. FJ178783.

Construct C, pET22b-EBNA1, the kind gift of Sarah Duellman, contains the wild-type EBNA1 nucleotide sequence (with amino-acids 92–315 deleted from the glycine-alanine region) cloned in expression vector pET22b [6].

Construct D, pET15b–$EBNA1_{DBD}$, contains the dimerization and DNA-binding domain of EBNA1 (amino-acids 452–641) cloned into pET15b.

Protein expression and purification

Escherichia coli strain BL21(DE3)pLysS or Rosetta 2 (Novagen) transfected with the plasmid of interest is grown in LB broth containing 100 μg/mL ampicillin, and induced at an OD (600 nm) of 0.6 with 1 mM IPTG for 4 h. Proteins are extracted as indicated in

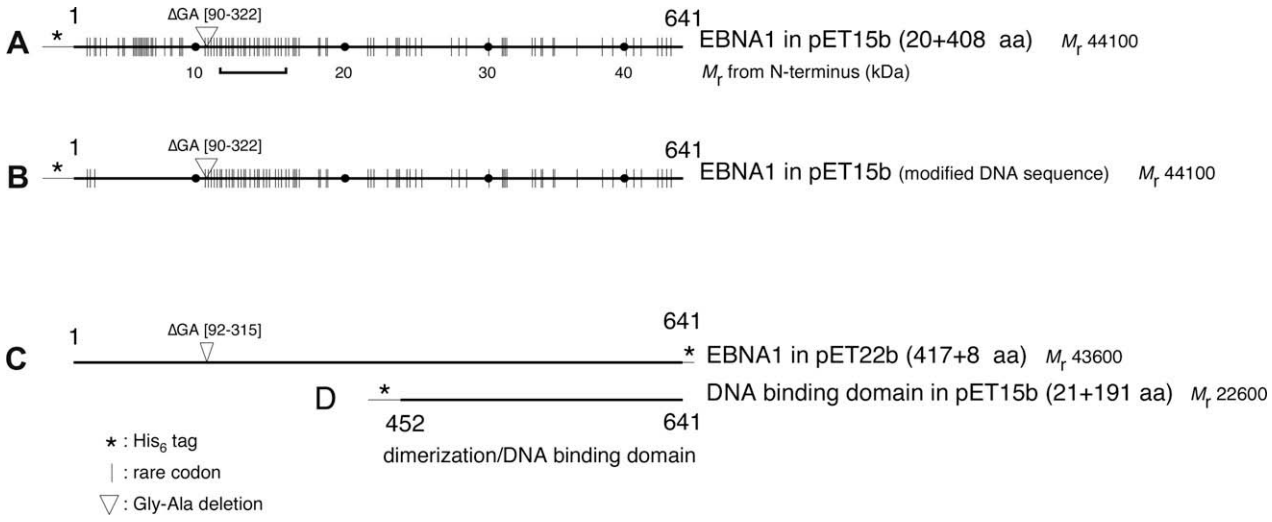

Fig. 1. Map of EBNA1 gene as inserted in plasmid constructs for EBNA1 expression. The first two constructs, A and B, contained the EBNA1 gene (with a deletion of the glycine and alanine domain as indicated) cloned into vector pET15b, for expression of a protein with a six-histidine tag at its N-terminus. Rare *E. coli* codons in constructs A and B are represented by vertical bars. Construct A contained the wild-type gene, construct B was derived from construct A by replacing all rare codons from position G_{15} to S_{78} by common *E. coli* codons coding for the same amino-acids. Construct C contained the EBNA1 gene cloned into vector pET22b [6] for expression of a protein with a C-terminal six-histidine tag. Construct D contained the part of EBNA1 gene coding for the C-terminal dimerization/DNA-binding domain cloned into vector pET15b. (*) indicate the location of the six-histidine tag in the expressed proteins. The horizontal bracket between 10 and 20 kDa under construct A indicates the C-terminal ends of the short EBNA1 polypeptides observed in Figs. 2 and 3.

the QIAexpressionist manual (Qiagen) for preparation of cleared *E. coli* lysates under native conditions, with the following modifications: the NaCl concentration in buffers is raised to 1 M through the whole extraction and purification process; 10 mM 2-mercaptoethanol is present in all buffers; protease inhibitors pepstatin 2 μg/mL, bestatin 10 μg/mL, chymostatin 10 μg/mL, and leupeptin 25 μg/mL are added at the lysozyme incubation step. After sonication and centrifugation, the cleared lysate is loaded onto Ni–NTA resin equilibrated with phosphate–imidazole buffer containing 1 M NaCl, and EBNA1 is eluted with increasing concentrations of imidazole, always in the presence of 1 M NaCl. Fractions are analyzed by SDS–polyacrylamide electrophoresis and coomassie blue staining, the fractions containing essentially pure EBNA1 are pooled, dialyzed against 0.3 M NaCl, 25 mM Tris–HCl pH 7.6, 1 mM EDTA, 1 mM DTT, 50% glycerol, and stored in small aliquots at −70 °C.

Electrophoretic mobility shift assay

The DNA fragment with four EBNA1-binding sites was a 179 bp DNA fragment encompassing the DS region of *OriP*, and extending from position 9013 to 9191 of EBV strain B95-8 sequence. It was prepared by PCR from plasmid PCEP4 (Invitrogen). DNA fragments were dephosphorylated and ^{32}P labeled with T4 polynucleotide kinase in the presence of [γ-^{32}P]-ATP. For formation of DNA–protein complexes, EBNA1 protein was gently mixed with ~5000 cpm of labeled DNA fragment in 25 μL of 50 mM NaCl, 25 mM Tris–HCl pH 7.5, 1 mM DTT, 1 mM EDTA, 0.1% Triton X-100, 100 μg/mL BSA, 2% glycerol, plus 12 μg/mL poly(dI-dC). After 30 min at 37 °C, samples were loaded on a 4% polyacrylamide gel (acrylamide:bis-acrylamide 29:1) in 0.5× TBE buffer that had been preelectrophoresed for 10 min. After electrophoresis for 3.5 h at 10 V/cm, the gel was dried on Whatman 3MM paper and autoradiographed.

Results

The wild-type EBNA1 gene (deleted of the glycine and alanine-coding region, see Materials and methods) was cloned into *E. coli* vector pET15b. Upon induction of T7 RNA polymerase by IPTG, BL21(DE3) bacteria containing this construct were expected to express EBNA1 fused to a N-terminal sequence of 20 amino-acids including a six-histidine tag for affinity purification on Ni–NTA columns. The expected molecular weight was ~44 kDa, and no such protein could be detected (Fig. 2A, left lanes). The analysis of *E. coli* extracts only showed a series of prominent bands corresponding to shorter polypeptides of apparent molecular weight 20–25 kDa, that were absent in non-induced cells and strongly expressed upon T7 RNA polymerase induction, could bind Ni–NTA resin, and were detected in Western blots using anti-EBNA1 antibodies (Fig. 2B). Therefore, these low molecular weight proteins were very likely partial EBNA1 polypeptides extending from the N-terminus.

As the nucleotide sequence of EBNA1 gene contains a very high density of codons that are rare in *E. coli* (Fig. 1), including a series of 11 rare codons in a row (AA 40–50) in the first glycine-arginine region, and as Duellman and Burgess were able to express EBNA1 in *E. coli* by using a strain that overexpresses tRNAs for rare codons, we considered that EBNA1 expression in *E. coli* might be significantly enhanced by replacing the rare codons with codons that are common in *E. coli*. Rare codons between positions 15–78 of

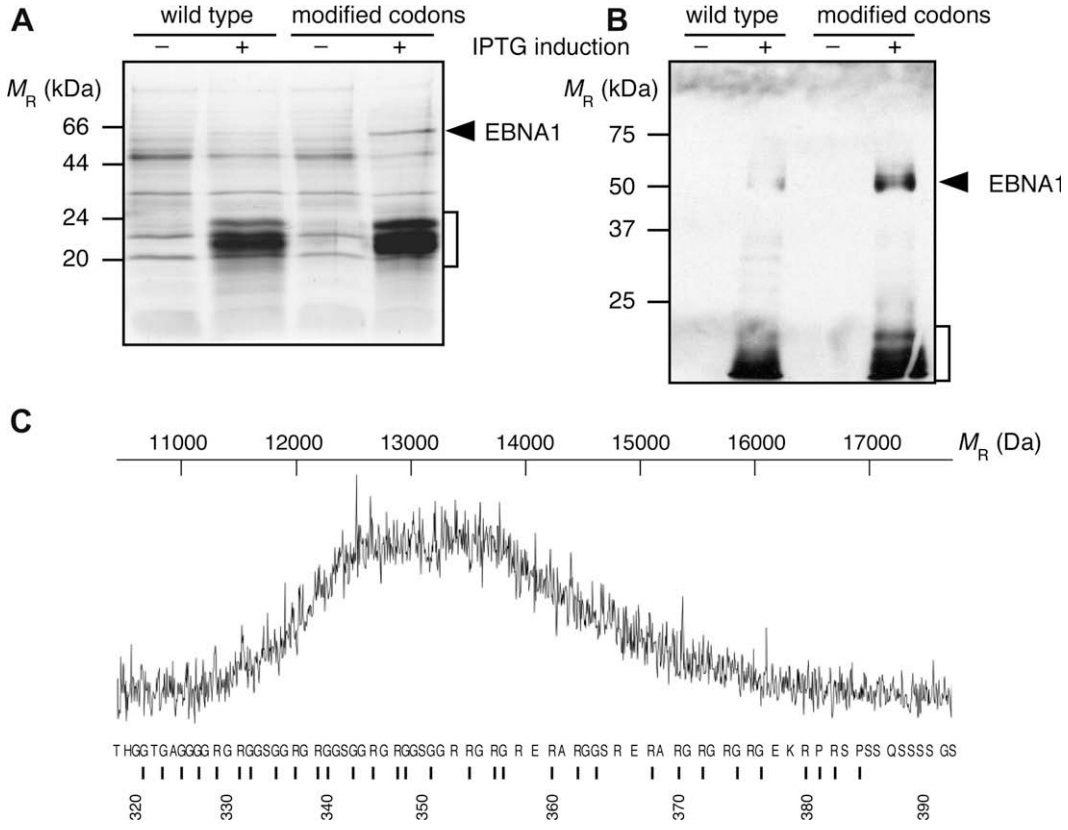

Fig. 2. Modification of rare codons of EBNA1 gene. EBNA1 was expressed in *E. coli* either from the wild-type nucleotide sequence, or with rare codons replaced with common codons as indicated in the text. The expressed proteins contained a N-terminal six-histidine tag. Total extracts from uninduced or induced bacteria were fractionated by affinity-chromatography on Ni–NTA resin, and the bound fractions were analyzed by electrophoresis on SDS–polyacrylamide gel. (A) Coomassie-blue staining of the gel. (B) Analysis of the gel by Western blotting using specific anti-EBNA1 polyclonal antibodies. (C) MALDI-TOF mass spectrometry analysis of the short EBNA1 polypeptides. Molecular weights are indicated on the top. The amino-acid sequence, positions of rare *E. coli* codons (vertical bars), and amino-acid numbering along EBNA1 sequence, are indicated on the bottom.

the nucleotide sequence of EBNA1 gene were thus changed and the modified gene was expressed in a standard *E. coli* strain. As can be seen in Fig. 2, the modification of the gene sequence did improve EBNA1 expression, since a protein with a molecular weight corresponding to the full-length EBNA1 and reacting with anti-EBNA1 antibodies was expressed upon induction by IPTG. However, the partial EBNA1 polypeptides of low molecular weight were still present. The molecular weight of these short polypeptides was accurately determined by MALDI-TOF mass spectrometry (Fig. 2C), and found to range between 11.5 and 16 kDa, much lower than their apparent molecular weight of 20–25 kDa on polyacrylamide–SDS gels. Assuming that these polypeptides contain the N-terminal His_6 tag since they were purified on Ni–NTA, their C-terminal end could be mapped within the second glycine-arginine region (AA 325–377, Fig. 2C) which also contains a high density of rare codons. However expressing the same construct in *E. coli* Rosetta 2 strain did not solve the problem of the short polypeptides (Fig. 3A). In summary, the modification of codons between positions 15–78 improved EBNA1 expression similarly to what was obtained using the *E. coli* Rosetta 2 strain to provide bacteria with tRNAs for rare codons [6]. For expression of EBNA1 we therefore used EBNA1 cloned into pET22b [6], a construct which was kindly given to us by Sarah Duellman and that we expressed in *E. coli* Rosetta 2 strain, with the further advantage of the six-histidine tag at the C-terminus, allowing us to get rid of the short polypeptides by Ni–NTA affinity purification (see below).

Planning to use EBNA1 protein for DNA-binding studies, we wished to extract and purify it without using potentially denaturing chemicals, such as guanidine or urea, which are often used to minimize the solubility problems that are frequently encountered with EBNA1. Using a high salt concentration, 1 M sodium chloride added to all buffers throughout the whole process (sonication, extract centrifugation, Ni–NTA affinity chromatography), allowed us to avoid solubility problems without ever resorting to denaturing agents. Only after the final purification on Ni–NTA was the salt concentration lowered to 300 mM by dialysis against a glycerol-containing buffer. With this pET22b-derived construct we did not observe the presence of short EBNA1 polypeptide chains after purification on Ni–NTA resin. This confirms the suggestion made above that these polypeptides were premature termination products extending from the N-terminus, the presence of the six-histidine tag at the C-terminal end of the protein expressed from pET22b allowing us to get rid of the incomplete EBNA1 polypeptide chains and to purify EBNA1 to near homogeneity by affinity chromatography on Ni–NTA, as shown on Fig. 3B.

The protein was assayed for its interactions with DNA fragments containing specific EBNA1-binding sites. A 179 bp DNA fragment containing the DS region of *OriP* was purified, 5′-end labeled with ^{32}P, and used in band-shift assay experiments. As frequently observed with EBNA1, such experiments again posed solubility problems, especially since it is was not possible to increase the salt concentration to high values without inhibiting EBNA1-binding to DNA [10]. However using high concentrations of the polyanion poly(dI-dC) strongly limited the solubility problems without interfering significantly with DNA–protein interactions. Fig. 4 shows that under such incubation conditions up to four retarded bands could be observed on polyacrylamide gels with the DNA fragment containing the DS region, depending on the amount of protein added (Fig. 4C), as expected for EBNA1 binding to its four specific sites. The fact that the first three bands C1, C2, C3 are subdivided into two or more bands is most probably due to the fact that EBNA1 bends DNA and that the overall structure of the complex varies slightly depending upon which of the four EBNA sites are occupied, as observed previously [11,12]. Under identical conditions no retarded bands were formed with a control DNA fragment containing no EBNA1 site (Fig. 4A) and a single retarded band was observed with a DNA fragment containing a single specific site (Fig. 4B). The complexes obtained with the C-terminal dimerization/DNA-binding domain of EBNA1 binding to the DS fragment are shown for comparison (Fig. 4D).

In summary, EBNA1 can be expressed in *E. coli* provided that the problem posed by the presence of rare codons in the nucleotide sequence is taken into account, but its expression is somewhat limited by additional factors that result in prematurely terminated polypeptide chains. EBNA1 can be extracted and purified from *E. coli* in soluble form when the salt concentration is kept high enough. The purified EBNA1 protein binds to DNA fragments containing specific EBNA1-binding sites and, under appropriate conditions, forms complexes that do not aggregate and can be analyzed by EMSA.

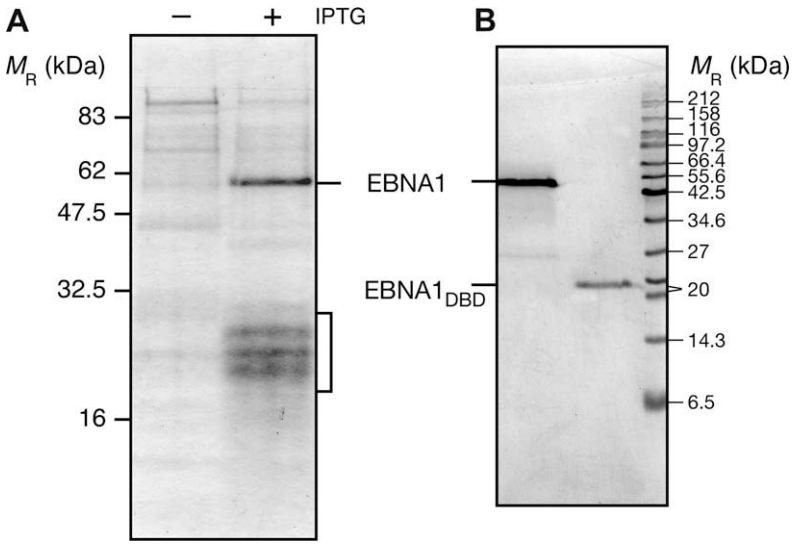

Fig. 3. Expression of EBNA1 in *E. coli* Rosetta 2 strain, which expresses tRNAs for rare codons. Wild-type EBNA1 gene was expressed as a fusion protein with a six-histidine tag either at its N-terminus (panel A) or C-terminus (panel B). Extracts prepared from bacteria without using denaturing agents (see Materials and methods) were fractionated on Ni–NTA, and the six-histidine tagged proteins were analyzed by SDS–polyacrylamide gel electrophoresis and Coomassie-blue staining. The C-terminal dimerization/DNA-binding domain of EBNA1 expressed with a N-terminal six-histidine tag is also shown on panel B, lane 2 (EBNA1$_{DBD}$).

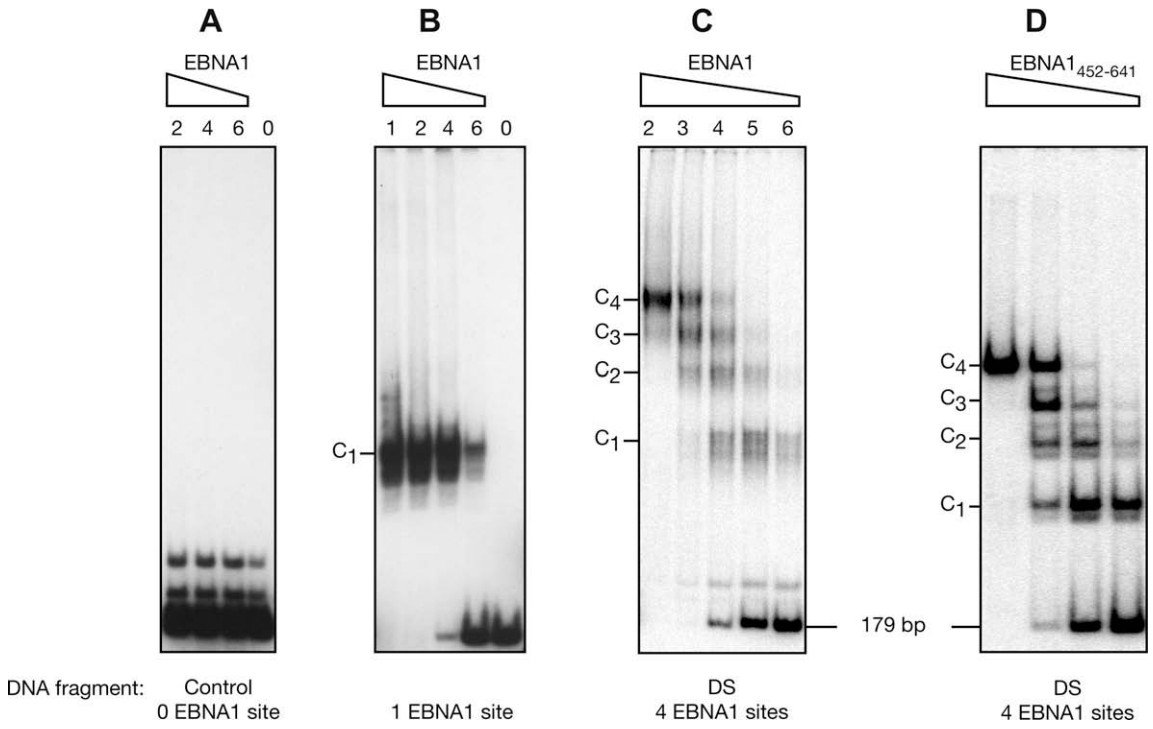

Fig. 4. Interactions of EBNA1 expressed in *E. coli* with DNA. EBNA1 protein, expressed in *E. coli* as a fusion with a C-terminal six-histidine tag and purified under non-denaturing conditions as described in the text, was analyzed by EMSA with ^{32}P-labeled DNA fragments containing 0, 1, or 4 EBNA1-binding sites in (A)–(C), respectively. Interactions of the C-terminal dimerization/DNA-binding domain of EBNA1 (residues 452–641) with the DNA fragment containing four binding sites are also shown for comparison (D). A given amount of labeled DNA (~200 pg per lane) was incubated with decreasing concentrations of EBNA1 (30, 10, 3, 1, 0.3, 0.1 ng in lanes labeled 1–6, respectively) in the conditions described in Materials and methods. DNA–protein complexes were analyzed by electrophoresis on a non-denaturing 4% polyacrylamide gel. Autoradiograms of the gels are shown. Weak bands above the free double-stranded DNA fragments in panels A and C are due to partial DNA denaturation.

Discussion

Full-length EBNA1 has been known to be difficult to express in *E. coli*. This is in contrast to the expression of the 190-aa C-terminal region that comprises the EBNA1 dimerization and DNA-binding domain. Previous biochemical studies on full-length EBNA1 have mostly been performed on protein expressed in insect cells. While this is a time-consuming and expensive system, it has the advantage of good expression levels and of producing a protein that contains eukaryotic post-translational modifications. It remains to be studied whether such modifications are similar in insect cells and in human cells, since the latter constitute the natural host of EBV. Since EBNA1 expressed in *E. coli* is devoid of eukaryotic modifications, it should provide an useful tool to study the difference between modified and unmodified proteins and to analyze the role of post-translational modifications for EBNA1 function in human cells. EBNA1 has no known enzymatic function, which should not be taken to imply that it has no enzymatic function at all: its exact biochemical role remains to be precisely defined and EBNA1 expressed in *E. coli* should be extremely useful for such studies.

Rare codons play an important role in the low level of expression of EBNA1, as shown by Duellman and Burgess [6] and confirmed by the present work. By changing all rare codons between amino-acid positions 15 and 78 we were able to express the full length protein without resorting to an *E. coli* strain engineered to express tRNAs for rare codons. We changed the rare codons in that region because of their particularly high density in the nucleotide sequence, with an extreme case of 11 rare codons in a row at positions 40–50. It might also be interesting to modify the rare codons in region 323–384, where they are still much more frequent than in the C-terminal region. It is unclear whether rare codons are responsible for the accumulation of short polypeptides ending at position 330–375. First, a similar accumulation of short polypeptides is observed when expression is performed in Rosetta 2 strain (Fig. 3A), suggesting that rare codons are not responsible for premature translation termination; second, short polypeptides do not terminate on the sides of this region, between 323–330 and 375–384, where rare codons are also frequent. The possibility remains, however, that expression of rare tRNAs by Rosetta 2 is not sufficient to overcome the high concentration of rare codons in this region. Another possibility is that the large number of G nucleotides in the mRNA, due to the richness in G of arginine and glycine codons, results in premature end of translation. When changing the rare codons in this region, it might thus be interesting to also modify the mRNA secondary structure and study whether this might have an effect on the efficiency of translation, as shown for the adjacent glycine–alanine repeat region where the introduction of mRNA secondary structures was found to enhance protein synthesis [13]. In this case, it would be interesting to study whether EBNA1 polypeptides ending in the arginine–glycine region are found among the defective ribosomal products (DRiPs) that have been described for EBNA1 in certain cell types [14]. In such a case, this glycine–arginine region would possibly play a role similar to that of its adjacent glycine-arginine region in the presentation of EBNA1 peptides to CD8(+) T cells ([15], recent review by [16]).

Using 1 M NaCl during EBNA1 extraction and purification is a very efficient way to avoid solubility and aggregation problems. Previously, high salt in combination with guanidine hydrochloride have been used for EBNA1 extraction, a concentration of only 0.5 M of GuHCl being very efficient to solubilize EBNA1 expressed in *E. coli* [6]. For most enzymes studied in the literature such a low concentration of guanidine does not induce denaturation (review [17]), and in the specific case of EBNA1 the DNA-binding activity

was fully preserved [6]. Examples have been reported, however, of enzymes irreversibly inactivated by low guanidine concentration (see e.g. [18,19], also references in [17]). Therefore, as the possibility cannot be excluded of a still-unknown enzymatic activity of EBNA1, a purification method avoiding potentially denaturing agents may eventually prove very useful.

The use of high concentrations of poly(dI-dC) to study the interactions of EBNA1 with its specific DNA-binding sites is particularly instructive, as it suppresses the aggregation ("linking") phenomenon and leads to the formation of well-defined DNA-EBNA1 complexes that migrate properly upon electrophoresis in polyacrylamide gels. Other polyanions gave similar results (manuscript in preparation), whereas omitting poly(dI-dC) from the incubation medium resulted in all labeled DNA remaining at the top of the gel (data not shown, see e.g. [20]).

The property of EBNA1 to link DNA fragments that contain its binding sites has been suggested to be one of the mechanisms by which EBNA1 plays a role in the segregation of viral episomes [21,22]. It is interesting to note that, in our experiments, linking is no longer observed in the presence of a concentration of DNA (containing no specific EBNA1-binding sites) that is much lower than the DNA concentration in the nucleus. Indeed, the DNA concentration is only ~10 μg/mL in our experiments, whereas it is about three orders of magnitude higher in the nucleus of a eukaryotic cell. By some aspects, our results strongly suggest that EBNA1 behaves very differently in the dilute conditions generally used for *in vitro* studies, and in the high-concentration conditions that are found *in vivo*, where molecular crowding is so high that it can strongly modify the apparent properties of macromolecules [23,24].

Acknowledgments

We are grateful to Sarah Duellman for the kind gift of the pET22b-EBNA1 expression vector. We also thank Gérard Bolbach, from the Mass Spectrometry and Proteomics Facility of Pierre et Marie Curie University (Paris), for the analysis of short EBNA1 peptides. C.G. and F.S. are members of the Centre National de la Recherche Scientifique (CNRS). N.B. was the recipient of a Ph.D. fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche.

References

[1] S.E. Lindner, B. Sugden, The plasmid replicon of Epstein-Barr virus: Mechanistic insights into efficient, licensed, extrachromosomal replication in human cells, Plasmid 58 (2007) 1–12.
[2] M. Conese, C. Auriche, F. Ascenzioni, Gene therapy progress and prospects: episomally maintained self-replicating systems, Gene Ther. 11 (2004) 1735–1741.
[3] D.A. Jackson, S. Juranek, H.J. Lipps, Designing nonviral vectors for efficient gene transfer and long-term gene expression, Mol. Ther. 14 (2006) 613–626.
[4] A. Bochkarev, J.A. Barwell, R.A. Pfuetzner, E. Bochkareva, L. Frappier, A.M. Edwards, Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein, EBNA1, bound to DNA, Cell 84 (1996) 791–800.
[5] A. Bochkarev, J.A. Barwell, R.A. Pfuetzner, W. Furey Jr., A.M. Edwards, L. Frappier, Crystal structure of the DNA-binding domain of the Epstein-Barr virus origin-binding protein EBNA 1, Cell 83 (1995) 39–46.
[6] S.J. Duellman, R.R. Burgess, Overproduction in *Escherichia coli* and purification of Epstein-Barr virus EBNA-1, Protein Expr. Purif. 47 (2006) 434–440.
[7] J.F. Kane, Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*, Curr. Opin. Biotechnol. 6 (1995) 494–500.
[8] C. Gustafsson, S. Govindarajan, J. Minshull, Codon bias and heterologous protein expression, Trends Biotechnol. 22 (2004) 346–353.
[9] S.J. Duellman, R.R. Burgess, Large-scale Epstein-Barr virus EBNA1 protein purification, Protein Expr. Purif. 63 (2009) 128–133.
[10] L. Frappier, M. O'Donnell, Overproduction, purification, and characterization of EBNA1, the origin binding protein of Epstein-Barr virus, J. Biol. Chem. 266 (1991) 7819–7826.
[11] J.L. Yates, S.M. Camiolo, J.M. Bashaw, The minimal replicator of Epstein-Barr virus *oriP*, J. Virol. 74 (2000) 4512–4522.
[12] J.M. Bashaw, J.L. Yates, Replication from *oriP* of Epstein-Barr virus requires exact spacing of two bound dimers of EBNA1 which bend DNA, J. Virol. 75 (2001) 10603–10611.
[13] J. Tellam, C. Smith, M. Rist, N. Webb, L. Cooper, T. Vuocolo, G. Connolly, D.C. Tscharke, M.P. Devoy, R. Khanna, Regulation of protein translation through mRNA structure influences MHC class I loading and T cell recognition, Proc. Natl. Acad. Sci. USA 105 (2008) 9319–9324.
[14] J. Tellam, G. Connolly, K.J. Green, J.J. Miles, D.J. Moss, S.R. Burrows, R. Khanna, Endogenous presentation of CD8+ T cell epitopes from Epstein-Barr virus-encoded nuclear antigen 1, J. Exp. Med. 199 (2004) 1421–1431.
[15] Y. Yin, B. Manoury, R. Fahraeus, Self-inhibition of synthesis and antigen presentation by Epstein-Barr virus-encoded EBNA1, Science 301 (2003) 1371–1374.
[16] M.E. Ressing, D. Horst, B.D. Griffin, J. Tellam, J. Zuo, R. Khanna, M. Rowe, E.J. Wiertz, Epstein-Barr virus evasion of CD8(+) and CD4(+) T cell immunity via concerted actions of multiple gene products, Semin. Cancer Biol. 18 (2008) 397–408.
[17] C.L. Tsou, Inactivation precedes overall molecular conformation changes during enzyme denaturation, Biochim. Biophys. Acta 1253 (1995) 151–162.
[18] A. Mitraki, J.M. Betton, M. Desmadril, J.M. Yon, Quasi-irreversibility in the unfolding–refolding transition of phosphoglycerate kinase induced by guanidine hydrochloride, Eur. J. Biochem. 163 (1987) 29–34.
[19] R.F. Jiang, Z.X. Wang, G.J. Xu, Substrate induced reactivation of spinach ribulose-1, 5-bisphosphate carboxylase/oxygenase denatured by low concentrations of guanidine hydrochloride, Biochim. Biophys. Acta 1343 (1997) 95–101.
[20] T.M. Avolio-Hunter, L. Frappier, Mechanistic studies on the DNA linking activity of Epstein-Barr nuclear antigen 1, Nucl. Acids Res. 26 (1998) 4462–4470.
[21] L. Frappier, M. O'Donnell, Epstein-Barr nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein-Barr virus, Proc. Natl. Acad. Sci. USA 88 (1991) 10875–10879.
[22] W. Su, T. Middleton, B. Sugden, H. Echols, DNA looping between the origin of replication of Epstein-Barr virus and its enhancer site: stabilization of an origin complex with Epstein-Barr nuclear antigen 1, Proc. Natl. Acad. Sci. USA 88 (1991) 10870–10874.
[23] R.J. Ellis, Macromolecular crowding: obvious but underappreciated, Trends Biochem. Sci. 26 (2001) 597–604.
[24] A.P. Minton, Macromolecular crowding, Curr. Biol. 16 (2006) R269–R271.