

**OVER-EXPRESSION, PURIFICATION AND CRYSTALLIZATION  
OF EPSTEIN-BARR NUCLEAR ANTIGEN-1**

OVER-EXPRESSION, PURIFICATION AND CRYSTALLIZATION  
OF THE DNA BINDING AND DIMERIZATION REGION OF  
EPSTEIN-BARR NUCLEAR ANTIGEN-1

by

Jean A. Barwell, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Sciences

McMaster University

April 1995

MASTER OF SCIENCE (1995)  
(Biochemistry)

McMASTER UNIVERSITY  
Hamilton, Ontario

TITLE: Over-expression, purification and  
crystallization of Epstein-Barr Nuclear  
Antigen-1.

AUTHOR: Jean A. Barwell, B.Sc. (University of Guelph)

SUPERVISOR Dr. Lori D. Frappier

NUMBER OF PAGES: xi, 55

## ABSTRACT

EBV episomes replicate once per cellular S phase, during latent infection of host cells. Only one viral protein, Epstein-Barr Nuclear Antigen-1 (EBNA-1) is required for replication; the rest of the replication machinery is provided by the cell. EBNA-1 is an excellent model to study the molecular events required for DNA replication and its regulation because viral replication is limited to once per cell cycle. EBNA-1 is a member of a special class of DNA binding proteins called origin binding proteins (OBPs). These specialized proteins bind to distinct DNA sequences in the genome called origins of replication, where DNA replication is initiated. Origin binding proteins may serve to distort the DNA at the origin and may also attract the cellular replication machinery. Structural studies of the DNA binding and dimerization region of EBNA-1 using X-ray crystallography were undertaken in order to better understand how OBPs bind to origin DNA sequences and facilitate the assembly of the cellular replication apparatus. Six truncation mutants of EBNA-1, all containing the DNA binding and dimerization region of EBNA-1, were cloned, over-expressed in bacteria and purified to apparent homogeneity. Four of these clones were crystallized using the method of hanging-drop vapour-diffusion. Two fragments, EBNA<sub>470-619</sub> and EBNA<sub>470-607</sub>, formed well-ordered crystals that diffracted beyond 2.5 Å



resolution. In addition, this study also demonstrates the value of finding the most suitable piece of the protein for crystallization. This piece should fold into a compact domain for efficient packing into a crystal. Finding the optimal piece of the protein reduces the time spent searching for crystallization conditions.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Aled Edwards and Dr. Lori Frappier for patiently teaching me so much about biochemistry and also for their contributions to this thesis. Technical assistance from Richard Pfuetzner was extremely helpful. Dr. Dan Yang's and Dr. Frank Sicheri's advice and assistance in the initial X-ray analysis of the crystals was invaluable. I would also like to thank Dr. Alexey Bochkarev for performing the X-ray analysis. Continuing friendships with everyone from the laboratories of Dr. Edwards and Dr. Frappier were both enjoyable and indispensable. I would finally like to acknowledge my family and my fiance, Vince Kloosterman, for their support and encouragement during my studies.

## TABLE OF CONTENTS

	Page
Abstract	iii
Acknowledgements	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
Abbreviations	x
I. INTRODUCTION	
I.1. Origin Binding Proteins	1
I.2. Latent EBV DNA replication	2
I.3. The EBV latent origin of replication	2
I.4. Epstein Barr Nuclear Antigen-1	3
I.4.i. DNA Binding by EBNA-1	4
I.4.ii. Transcriptional enhancement by EBNA-1	5
I.4.iii. EBNA-1/DNA complex formation at <i>oriP</i>	6
I.4.iv. Mutagenesis of EBNA-1	7
I.5. Protein interactions at <i>oriP</i>	8

I.6. Activation of replication origins by OBPs	9
I.7. Formation of replication initiation complex	10
I.8. Structural analysis of EBNA-1	12
II. MATERIALS AND METHODS	
II.1. Cloning of the EBNA-1 truncation mutants	14
II.2. Overexpression of the EBNA-1 truncation mutants	15
II.3. Purification of the EBNA-1 truncation mutants	16
II.4. Proteolysis of the EBNA-1 truncation mutants	22
II.5. Electrophoretic mobility shift assays	22
II.6. Preparation of radioactive dyad symmetry element	23
II.7. Crystallization	23
II.8. Data Collection	24
III. RESULTS	
III.1. Expression of the EBNA-1 truncation mutants	25
III.2. Purification	27
III.3. The search for the most suitable fragment of EBNA-1 for crystallization	28
III.4. Protein crystallization	33
IV. DISCUSSION	44
REFERENCES	49

## LIST OF FIGURES

Figure		Page
1a	Purification scheme of EBNA <sub>452-641</sub> and EBNA <sub>459-619</sub>	19
1b	Purification scheme of EBNA <sub>470-607</sub> and EBNA <sub>468-607</sub>	20
1c	Purification scheme of EBNA <sub>459-607</sub>	21
2	EBNA-1 truncation mutants	26
3	SDS polyacrylamide gel of the EBNA-1 truncation mutants	31
4	Mass spectroscopy spectra of EBNA <sub>470-607</sub>	32
5	Mobility shift assay of EBNA-1 truncation mutants	36
6a	Crystals of EBNA <sub>452-641</sub> grown in ammonium phosphate	37
6b	Crystals of EBNA <sub>452-641</sub> grown in ammonium phosphate	38
6c	Crystals of EBNA <sub>452-641</sub> grown on lipid layers	39
6d	Crystals of EBNA <sub>470-619</sub> grown in sodium acetate	40
6e	Crystals of EBNA <sub>470-619</sub> grown in PEG 4K	41
6f	Crystals of EBNA <sub>470-619</sub> grown in sodium formate	42
6g	Crystals of EBNA <sub>470-607</sub> grown in NaCl	43

## LIST OF TABLES

Table	Page
Purification of EBNA-1 truncation mutants	30



## ABBREVIATIONS

Å	Ångstrom
ATP	Adenosine tri-phosphate
CAT	chloramphenicol acyltransferase
cDNA	complementary DNA
DMS	dimethylsulfate
DNA	deoxyribonucleic acid
DS	dyad symmetry
DTT	dithiothreitol
EBNA-1	Epstein-Barr Nuclear Antigen-1
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
FR	family of repeats
HSV	Herpes simplex virus
kb	kilobase
L	litre
M	molar
mM	millimolar

mg	milligram
ml	millilitre
NaAc	sodium acetate
OBP	origin binding protein
O.D.	optical density
pCMB	parachloromercuribenzoate
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
RNA	ribonucleic acid
RPA	replication protein A
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SV40	simian virus 40
μg	microgram

## **I. INTRODUCTION**

### **I.1. Origin Binding Proteins**

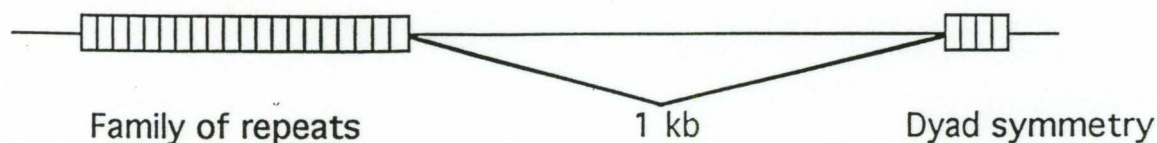
Initiation of DNA replication occurs at specific sites throughout the genome called origins of replication. When these distinct DNA sequences are bound by specialized DNA binding proteins called origin binding proteins (OBPs), a sequence of molecular events leading to the initiation of DNA replication occurs (reviewed in Stillman, 1994). While their modes of action are being studied in many organisms including bacteria, phages, animal DNA viruses and yeast, the mechanism by which the OBPs initiate replication is not known. OBPs across species have no homology nor do they have any sequence homology to any characterized DNA binding protein. However, one common feature of all OBPs is the formation of multimeric protein-DNA complexes at the origin. The OBPs and the formation of these complexes may serve to unwind the DNA or make the DNA more accessible to DNA helicases. This complex and the distorted DNA may also act to attract the cellular replication machinery.

## I.2. Latent EBV DNA replication

During latent infection of host cells, the EBV genome is maintained as a circular episome in the cell (Shaw et al., 1979). Bi-directional replication occurs autonomously from the latent origin of replication once per cellular S phase (Adams, 1987; Yates and Guan, 1991). The episomes segregate stably over many cell divisions (Yates et al., 1984, 1985). EBNA-1 is the only viral-encoded protein required for this replication; all other replicative functions are provided by the host cell. EBNA-1 serves as an excellent model to better understand the molecular events necessary for DNA replication and its regulation in mammalian systems, because replication initiated at the latent origin of replication is limited to once every cellular S phase during latent infection (Adams, 1987; Yates and Guan, 1991).

## I.3. The EBV latent origin of replication

### OriP



Yates et al. (1984) discovered the EBV latent origin of



replication, *oriP*, as a *cis*-acting sequence allowing the replication and maintenance of plasmids in cells latently infected with EBV. *OriP* is composed of two elements, the family of repeats (FR) and the dyad symmetry element (DS) (Reisman et al., 1985). The FR is composed of 20 tandem copies of a 30 base pair repeat. Within each repeat is the 18 base pair palindromic EBNA-1 binding site. The DS contains four EBNA-1 binding sites. The FR and DS are separated by 960 base pairs of DNA which can be deleted with no effect on *oriP* function (Reisman et al., 1985). Two-dimensional gel electrophoresis has revealed that DNA replication initiates at or very near to the dyad symmetry element (Gahn and Schildkraut, 1989).

#### **I.4. Epstein Barr Nuclear Antigen-1**

EBNA-1 has been shown to play a role in stable plasmid segregation, viral DNA replication (Yates et al., 1985), and transcriptional enhancement (Reisman and Sugden, 1986). While the mode of action of EBNA-1 is not understood for these functions, the DNA binding and dimerization region of the protein is required. EBNA-1 binds as a dimer to the binding sites within the FR and DS (Frappier and O'Donnell, 1991; Rawlins et al., 1985; Ambinder et al., 1991).

#### **I.4.i. DNA Binding by EBNA-1**

The ability of EBNA-1 to bind DNA is central to all functions of the protein. One role EBNA-1 may serve is to distort the DNA at the origin of replication perhaps making it more accessible to cellular helicases. DNA unwinding is necessary for the initiation of DNA replication (DePamphilis, 1993) and the formation of the replication bubble. While EBNA-1 has been shown to have no helicase activity (Frappier and O'Donnell, 1991) or ability to melt DNA (Frappier and O'Donnell, 1992), it has been shown to distort or bend the DNA at the dyad symmetry element. Two thymine residues in the dyad symmetry element were shown to be reactive to  $\text{KMnO}_4$ .  $\text{KMnO}_4$  selectively oxidizes pyrimidines in single-stranded or helically distorted DNA. Binding of EBNA-1 to the DS element also causes nucleotides on one face of the helix to be protected from DMS methylation while nucleotides on the opposite face are hypersensitive to DMS, suggesting that EBNA-1 bends the DNA towards itself (Frappier and O'Donnell, 1992). Distortion of DNA in the family of repeats is not detected (Hearing et al., 1992), suggesting that the differences in the spacing between binding sites in the FR and DS elements and the resulting protein-protein interactions between adjacent sites, was the cause of the difference in the ability to distort the DNA. Another explanation may be that the hyper-reactive thymine detected by sensitivity to  $\text{KMnO}_4$  is not present in any other EBNA-1 binding site. Distortion in these other binding



sites may therefore be undetectable. A recent study by Harrison et al. (1994) placed 5 or 10 base pairs between the binding sites in the DS element and found that while the correct spacing between sites in the DS element is essential for replicative ability, the DNA was still reactive to  $\text{KMnO}_4$ , indicating DNA distortion. Some protein-protein interaction may still have occurred on the mutated DS element, causing the DNA distortion, but the replicative ability was eliminated. Therefore, DNA distortion is likely necessary but not sufficient for replicative function.

Regulation of the initiation of replication has been thought to be controlled by EBNA-1 via its ability to cause distortion of the DNA at *oriP*. Hsieh et al. (1993) addressed this theory by performing *in vivo* footprints with DMS and  $\text{KMnO}_4$  and demonstrated that EBNA-1 is bound to the origin throughout the cell cycle and the DNA is constitutively held in a distorted conformation. Initiation at *oriP* is therefore likely under the control of a cellular mechanism and dependent on a cellular helicase.

#### **I.4.ii. Transcriptional enhancement by EBNA-1**

Along with the role of EBNA-1 in replication, this protein has also been shown to be active as a transcriptional enhancer. When EBNA-1 is provided *in trans*, the family of repeats, when positioned either upstream or downstream of the chloramphenicol

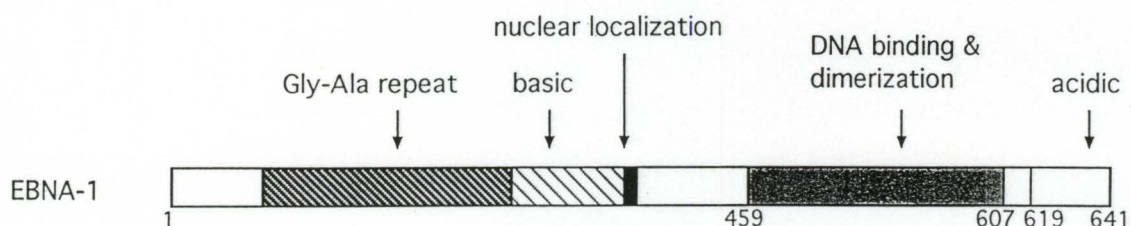
acyltransferase (CAT) gene, can act as an enhancer of expression from either the SV40 early promoter or the HSV-1 thymine kinase promoter (Reisman and Sugden, 1986). The interaction of EBNA-1 with the family of repeats also enhances transcription from the Cp promoter on the BamHI-C fragment of EBV, which is active during latent EBV infection (Sugden and Warren, 1989). This promoter has been shown to drive EBNA-1 expression in some cases which suggests that EBNA-1 may affect its own expression. In another study, EBNA-1 was shown to negatively affect expression from the Fp promoter which is the only promoter used for EBNA-1 expression during type I latency, indicating that EBNA-1 can negatively regulate its own expression (Sample et al., 1992).

#### **I.4.iii. EBNA-1/DNA Complex Formation at *OriP***

When EBNA-1 binds to the family of repeats and the dyad symmetry element within *oriP*, these two regions of *oriP* interact via an EBNA-1 complex, causing looping out of the 1 kb of intervening DNA (Frappier and O'Donnell, 1991; Su et al., 1991). This complex formation has been found to stabilize the interaction of EBNA-1 with the dyad symmetry element (Su et al., 1991, Frappier et al., 1994). Adjacent EBNA-1 molecules at the dyad symmetry element also interact (Hearing et al., 1992; Harrison et al., 1994; Summers et al., *manuscript in preparation*), perhaps further stabilizing binding to the

origin. An important role of EBNA-1 may be to stabilize assembly or facilitate activity of the replication initiation complex.

#### I.4.iv. Mutagenesis of EBNA-1



Studies of deletion mutants of EBNA-1 has provided much of the information about functions of different regions of the polypeptide. The first third of the protein contains a repetitive sequence of glycines and alanines. This region can delete spontaneously and is dispensable for the replicative, transcription enhancement and plasmid maintenance functions of EBNA-1 (Yates et al., 1985). C-terminal to the Gly-Ala repeat, there is a basic domain of 56 amino acids and a nuclear localization signal from amino acid 379 to 386 (Ambinder et al., 1991). The region of the protein required for the DNA looping interaction between EBNA-1 bound to the family of repeats and EBNA-1 bound to the dyad symmetry element has been located between amino acids 350 and 362 (Frappier et al., 1994).

The region of interest to this study, the DNA



binding/dimerization domain, has been localized to amino acids 459-607 (Shah et al., 1992). Chen et al., (1993) further separated this region into the dimerization domain between amino acids 501 and 532 and amino acids 554 and 598, while the DNA binding region mapped between amino acids 459 and 487. Sequences in the region of amino acid 57 may also have some effect on DNA binding, perhaps by affecting protein stability (Polvino-Bodnar and Schaffer, 1992). The transactivation function of EBNA-1 has been difficult to localize. DNA binding is required for this function (Polvino-Bodnar and Schaffer, 1992). The acidic C-terminus of the protein is not required for transactivation by EBNA-1, but it can contribute to transactivation when expressed as a pyruvate kinase fusion protein (Ambinder et al., 1991). No mutation has been found which separates the replication and transactivation functions of EBNA-1 (Polvino-Bodnar et al., 1988; Yates and Camiolo, 1988).

### **I.5. Protein interactions at *OriP***

Attempts to find proteins which interact with EBNA-1 at the origin are on-going. The replicative cellular helicase which acts on *oriP* has not yet been found. EBNA-1 may interact with the single-stranded binding protein, RP-A (unpublished results) and evidence of *oriP* attachment to the nuclear matrix suggests interaction with cellular replication apparatus (Jankelevich et al., 1992). Some groups

have attempted to find cellular proteins which interact with the origin itself. Oh and Levine (1991) did footprinting experiments on *oriP* with Hela cell nuclear extracts. Five discrete footprints were detected, four of which overlapped with the EBNA-1 binding sites. Gel shift assays demonstrated that EBNA-1 could compete for binding to these sites. The authors proposed the existence of cellular factors which regulate viral DNA replication by excluding EBNA-1 from the *oriP* binding sites (Oh and Levine, 1991). Zhang and Nonoyama (1994) also report evidence of cellular factors which bind to *oriP* when the host cells were exposed to a tumour-promoting phorbol ester. A cDNA library was made from these cells and screened for proteins which bind to *oriP* and two c-DNAs were isolated.

#### **1.6. Activation of Replication Origins by OBPs**

As described above, one of the possible roles of EBNA-1 is activation of the origin via DNA distortion to allow formation of the replication bubble. Other origin binding proteins have also been shown to perform this role. In the presence of ATP, monomers of the Simian virus 40 (SV40) large tumour antigen (T-antigen) form a hexameric complex around the core *ori* sequence. DNA becomes distorted in regions flanking the core element, then *ori* becomes completely denatured while T-antigen unwinds the DNA in a bi-directional manner (reviewed in Borowiec et al., 1990). Another

origin binding protein,  $\lambda$  O protein of bacteriophage  $\lambda$ , also serves to distort the DNA at the phage origin,  $ori\lambda$ . O protein binds to four O protein binding sites and forms a protein/DNA complex. DNA winding around this complex causes distortions in the DNA.  $\lambda$  P protein then binds to the *E. coli* helicase, DnaB and brings the helicase to the origin by binding to  $\lambda$  O. *E. coli* heat shock proteins then inactivate  $\lambda$  P thereby allowing the helicase DnaB to unwind the DNA (reviewed in Stillman, 1994). Lastly, the herpes simplex virus (HSV) origin binding protein, UL9 binds to *oris* and loops and distorts the DNA and does have some helicase activity when stimulated by an HSV-encoded single-stranded binding protein, ICP8 (Koff et al., 1991; reviewed in DePamphilis, 1993).

### **I.7. Formation of Replication Initiation Complex**

Besides the role in DNA unwinding or distortion at the origin, OBPs are also thought to act as a "landing pad" (Stillman et al., 1994) for other replication proteins. For example, the  $\lambda$  O protein, when bound to the phage origin, binds to the phage P protein which is in turn bound to the *E. coli* DnaB helicase. Once replication is initiated by this complex of proteins, the primase and polymerase and accessory factors begin DNA synthesis (reviewed in Stillman, 1994).

The OBP of SV40, T-antigen, also forms the foundation of the replication complex. A double hexamer of T-antigen binds at the



origin in the presence of ATP and causes DNA distortion. This complex binds cellular replication protein A (RPA) which is a single-stranded DNA binding protein. This allows the T-antigen helicase function to unwind the DNA. The T-antigen-RPA complex binds polymerase- $\alpha$  ( $\text{pol}\alpha$ )/primase which synthesizes a nascent RNA-DNA at the origin (Waga and Stillman, 1994; reviewed in Borowiec et al., 1990).

One class of proteins which seems to be consistently involved with the activities of origin binding proteins and the replication machinery in both prokaryotes and eukaryotes is transcription factors (reviewed in DePamphilis, 1993). According to DePamphilis' model (1993), the origin of replication contains a core region where the OBP binds and auxiliary binding sites which bind transcription factors. Transcription factors may facilitate assembly of the initiation complex (DePamphilis, 1993). For example, the transcription factor, E2, and the origin binding protein, E1, of bovine papilloma virus (BPV) are both necessary and sufficient for viral replication (Yang et al., 1991). The origin core contains the binding site for E1 and it is flanked by two auxiliary binding sites for E2. The two proteins form a tight complex and E1 provides the helicase activity while E2 facilitates binding of E1 to the origin. In the presence of E2, the affinity of E1 for ori core is increased 10-fold.

SV40 T-antigen is an example of how transcription factors can facilitate the activity of the initiation complex once the complex has formed. In this case, again, ori core is flanked by two auxiliary

binding sites. When T-ag binds to ori core, it unwinds the DNA thereby destroying its own binding site. T-ag has low affinity for single-stranded DNA, therefore, in order to stabilize the unwound intermediate, T-ag binds to transcription factors bound to the auxiliary sites or T-ag itself binds to the auxiliary sites in order to stabilize the intermediate (Gutierrez et al., 1990; reviewed in DePamphilis, 1993).

As a transcription factor at the origin of replication, EBNA-1 bound to the family of repeats may serve to stabilize the EBNA-1/origin DNA complex through the DNA looping mechanism. While the family of repeats is not absolutely required for the molecular events leading to DNA replication initiation from the dyad symmetry element (Harrison et al., 1994), the formation of this complex may be necessary for long-term plasmid maintenance or regulation of origin activity (Harrison et al., 1994). The DNA looping mechanism may also serve to prevent chromatin structure from inhibiting formation of the replication complex (DePamphilis, 1993).

### **I.8. Structural Analysis of EBNA-1**

Elucidation of the molecular structure of EBNA-1 should provide information about the mechanism by which OBPs recognize origin DNA sequences and facilitate the assembly of the cellular replication apparatus. I have purified and crystallized the DNA

binding/dimerization domain of EBNA-1 as a first step towards this goal. Since DNA binding is essential for all known activities of EBNA-1, knowledge of this region should lend understanding to the mechanism of action of EBNA-1. By solving the structure of EBNA-1, we expect to learn how EBNA-1 performs the known functions of origin binding proteins, in particular, its effect on DNA bending, twisting or unwinding, its ability to bind other proteins involved in replication and the region of the protein required for transactivation.

I have chosen to redefine the borders of the DNA binding/dimerization domain in an effort to find the most suitable piece of the protein for crystallization (i.e. a piece of the protein which folds into a tight domain, suitable for packing in a crystal). I have approached this by limited proteolytic digestion and recloning the newly defined DNA binding domain for expression in bacteria. I purified several candidate truncation mutants of EBNA-1, explored many crystallization conditions and eventually found a suitable crystal form for X-ray diffraction analysis.



## II. MATERIALS AND METHODS

### II.1. Cloning of the EBNA-1 truncation mutants

I have cloned and purified six truncation mutants of EBNA-1, indicated below where the subscripts represent the amino acid coordinates within the protein. The coding sequence for each mutant was amplified from the EBNA1 gene in plasmid p205 (Yates et al, 1985) using Vent polymerase (New England Biolabs) by the polymerase chain reaction (PCR). The PCR primers were designed to yield amplification products with an NdeI site at the 5' end and a BamHI site at the 3' end and are as follows (written 5' to 3'):

EBNA<sub>452-641</sub>: 5' primer CGTCGACATATGGGTCAGGGTGATGGAGGC

3' primer CGTGCAGGATCCTCACTCCTGCCCTTCCTCACC

EBNA<sub>459-619</sub>: 5' primer CGTCGACATATGCGCAAAAAAGGAGGGTGG

3' primer CGTGCAGGATAATCACTCCGCGGCAGCCCCTTCCAC

EBNA<sub>470-619</sub>: 5' primer CCTCCACATATGGGTCAAGGAGGTTCCAAC

3' primer CGTGCAGGATAATCACTCCGCGGCAGCCCCTTCCAC

EBNA<sub>470-607</sub>: 5' primer CCTCCACATATGGGTCAAGGAGGTTCCAAC

3' primer CGTGCAGGATCCTCAAGGCAAATCTACTCCATC

EBNA<sub>468-607</sub>: 5' primer CGTCGACATATGCATCGTGGTCAAGGAGGT

3' primer CGTGCAGGATCCTCAAGGCAAATCTACTCCATC

EBNA<sub>459-607</sub>: 5' primer CGTCGACATATGCGCAAAAAACCACCCTGG

3' primer CGTCGAGGATCCTCAAGGCAAATCTACTCCATC

PCR products were purified by GeneClean (Bio101 Inc.) and the purified fragments were then digested for two hours with 20 units of NdeI (New England Biolabs) at 37°C. The NaCl concentration in the digest was then adjusted to 100 mM, and the DNA was digested for two hours with 20 units of BamHI (New England Biolabs) at 37°C. The digested DNA was then purified by the GeneClean method. The vector, pET15b (Novagen) was digested with NdeI and BamHI by the same procedure. Ligations were done at 15°C using 800 units of ligase (New England Biolabs) with 25 fmol of NdeI/BamHI digested pET15b and 100 fmol of NdeI/BamHI digested PCR product. Competent HB101 cells (100 ml) were transformed with half of the ligation reaction and plated on LB plates (10 g/L bactotryptone, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar) (Sambrook et al., 1989) containing 100 µg/ml ampicillin and incubated overnight at 37°C. Candidate clones were screened by mini-preparations (Sambrook et al., 1989) of the DNA and digestion with HindIII and XbaI (New England Biolabs).

## **II.2. Overexpression of the EBNA-1 truncation mutants**

The clones were transformed into 100 ml of *Escherichia coli* (BL21(DE3)pLysS) (Studier et al, 1990) (Novagen), plated on LB plates containing 100 mg/ml ampicillin and 34 mg/ml chloramphenicol and incubated overnight at 37°C. A culture (500 ml LB) was grown overnight at 37°C without shaking to an optical

density (O.D.) of 0.8, and used to inoculate 30 L of Terrific Broth (12 g/L bactotryptone, 24 g/L yeast extract, 4 ml/L glycerol, 0.017 M  $\text{KH}_2\text{PO}_4$ , 0.074 M  $\text{K}_2\text{PO}_4$ ) (Sambrook et al., 1989). Cells were grown in a fermentor (New Brunswick Scientific) with agitation (150 rpm), air flow at 12.5 litres per minute and air pressure at 30 psi. Once the culture reached an absorbance of 0.8 at 595 nm, EBNA-1 expression was induced with 0.5 mM isopropyl-B-D-thiogalactopyranoside, followed 30 minutes later by the addition of 150 mg/ml of rifampicin. Three hours post-induction, the cells were harvested by centrifugation and resuspended in 20mM Tris pH 7.5, 10% sucrose, 1.0 mM benzamidine, 1.0 mM phenylmethylsulfonyl fluoride, and 1.0 mM EDTA (1 ml per gram of cells).

### **II.3. Purification of the EBNA-1 truncation mutants**

Cells were frozen at  $-70^\circ\text{C}$  and then thawed and lysed by sonication. NaCl was added to 350 mM and, after a 30 minute incubation on ice, the lysate was clarified by centrifugation at 20,000 x g for 30 minutes. The supernatant was then placed in a  $75^\circ\text{C}$  water bath (for EBNA<sub>452-641</sub>) or in a  $55^\circ\text{C}$  water bath (for the other mutants) until the temperature of the protein solution reached the temperature of the bath. Heating was followed by a 10 minute incubation on ice. The supernatant was clarified by centrifugation at 20,000 x g for 30 min, and then flowed through a 60 ml DE52 column



(Whatman) equilibrated with 20 mM Tris pH 7.5, 350 mM NaCl, 10 mM dithiothreitol (DTT), and 10% glycerol. The DE52 flow-through was diluted with 50 mM Hepes pH 7.2 to a final NaCl concentration of 200 mM (for EBNA<sub>470-607</sub>, EBNA<sub>468-607</sub>, EBNA<sub>459-607</sub>) and loaded on to a 120 ml heparin-agarose column (Biorad) equilibrated with buffer A (50 mM Hepes pH 7.2, 200 mM NaCl, 10% glycerol, 10 mM DTT). EBNA<sub>452-641</sub>, EBNA<sub>459-619</sub> and EBNA<sub>470-619</sub> were loaded on to the heparin-agarose column in buffer A containing 350 mM NaCl. After washing the column with 4 column volumes of buffer A containing either 350 mM NaCl or 200 mM NaCl, the protein was eluted with buffer A containing 750 mM NaCl. The protein eluate was dialyzed against buffer B (50 mM Hepes pH 7.2, 750 mM NaCl, 10% glycerol, and 1 mM PMSF, 1 mM benzamidine), and loaded onto a 5 ml HPLC metal-chelating column (Perseptive Biosystems) charged with nickel and equilibrated in buffer B plus 5 mM imidazole. The column was developed with a 20 ml linear gradient from 5 mM to 300 mM imidazole in buffer B. The fractions containing >99% pure EBNA-1 (as judged by Coomassie staining of 15% SDS polyacrylamide gels) were pooled and digested with thrombin (10 units/mg of protein) at 37°C for 1 hour in buffer C (50 mM Hepes pH 7.2, 500 mM NaCl, 10 mM DTT, 10% glycerol, 2.5 mM CaCl<sub>2</sub>) for EBNA<sub>452-641</sub> and EBNA<sub>459-619</sub> and buffer C containing 750 mM NaCl for the other truncation mutants. Another 5 units of thrombin per mg of EBNA-1 was added and incubated for another 2 hours. The completion of the thrombin digestion was confirmed by Coomassie staining of the

protein on a 15% SDS polyacrylamide gel. After digestion, the protein was diluted to 1 mg/ml with the digestion buffer and then diluted further with buffer A lacking NaCl to a final NaCl concentration of 200 mM. This dilution protocol was necessary to avoid protein aggregation. The protein was loaded onto a 5 ml HPLC Poros S column (Perseptive Biosystems) at 200 mM NaCl in Buffer A, eluted with 1 M NaCl in buffer A at a slow flow rate (1 ml/min), and concentrated to 10 mg/ml using a Centricon-10 micro-concentrator (Amicon) if necessary. The purified protein was finally dialyzed into 1 mM Hepes pH 7.2, 10 mM DTT, and 500 mM NaCl (buffer C) for EBNA<sub>452-641</sub>, EBNA<sub>459-619</sub> and EBNA<sub>470-619</sub> and buffer C containing 750 mM NaCl for EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub>, and buffer C containing 1.0 M NaCl for EBNA<sub>459-607</sub> (see Figure1 a,b,c).

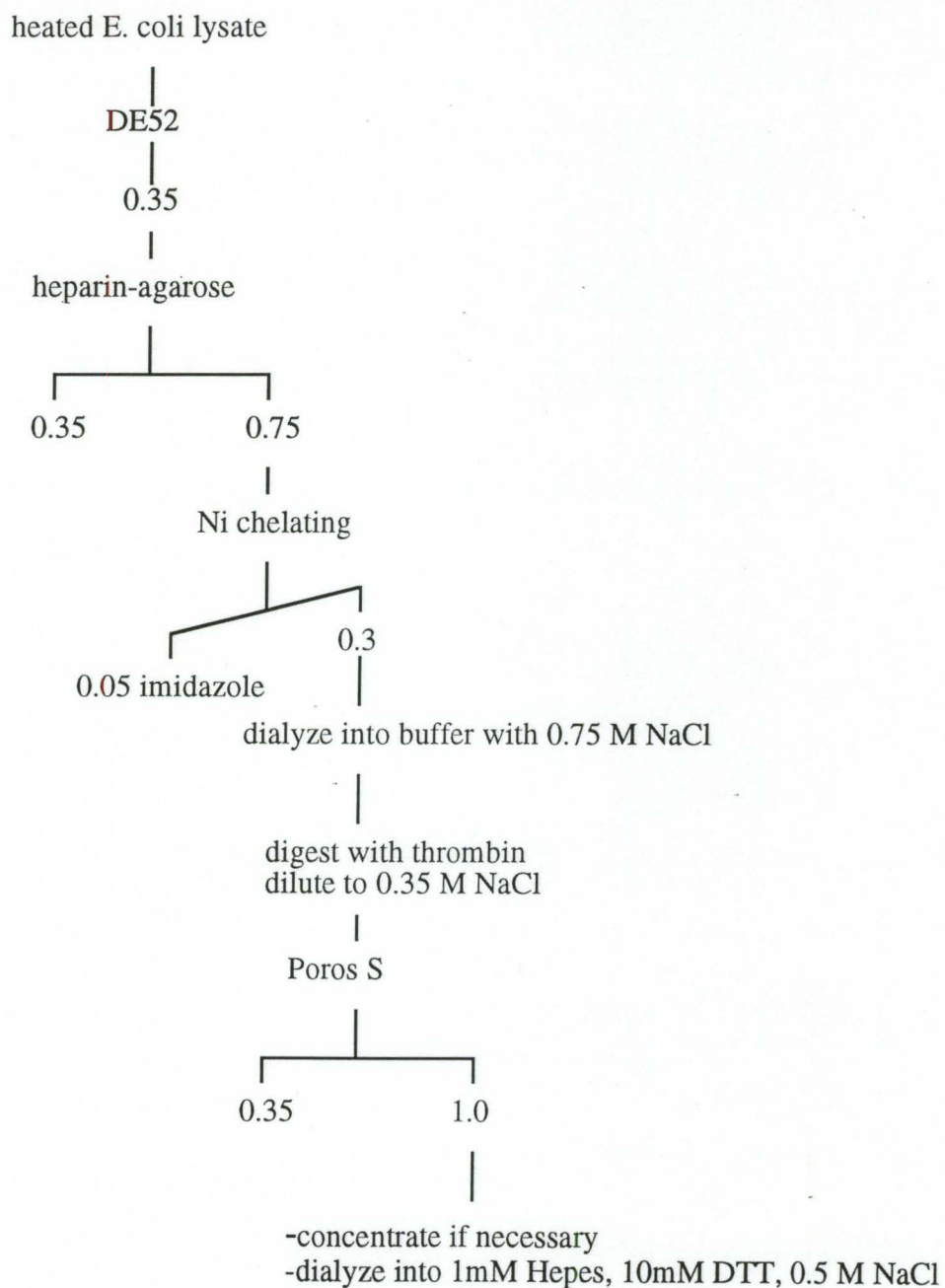


Figure 1a. Purification scheme of EBNA<sub>452-641</sub> and EBNA<sub>459-619</sub>. Horizontal and diagonal lines represent step and gradient elutions, respectively. Elution was with sodium chloride with concentrations given in moles/litre.

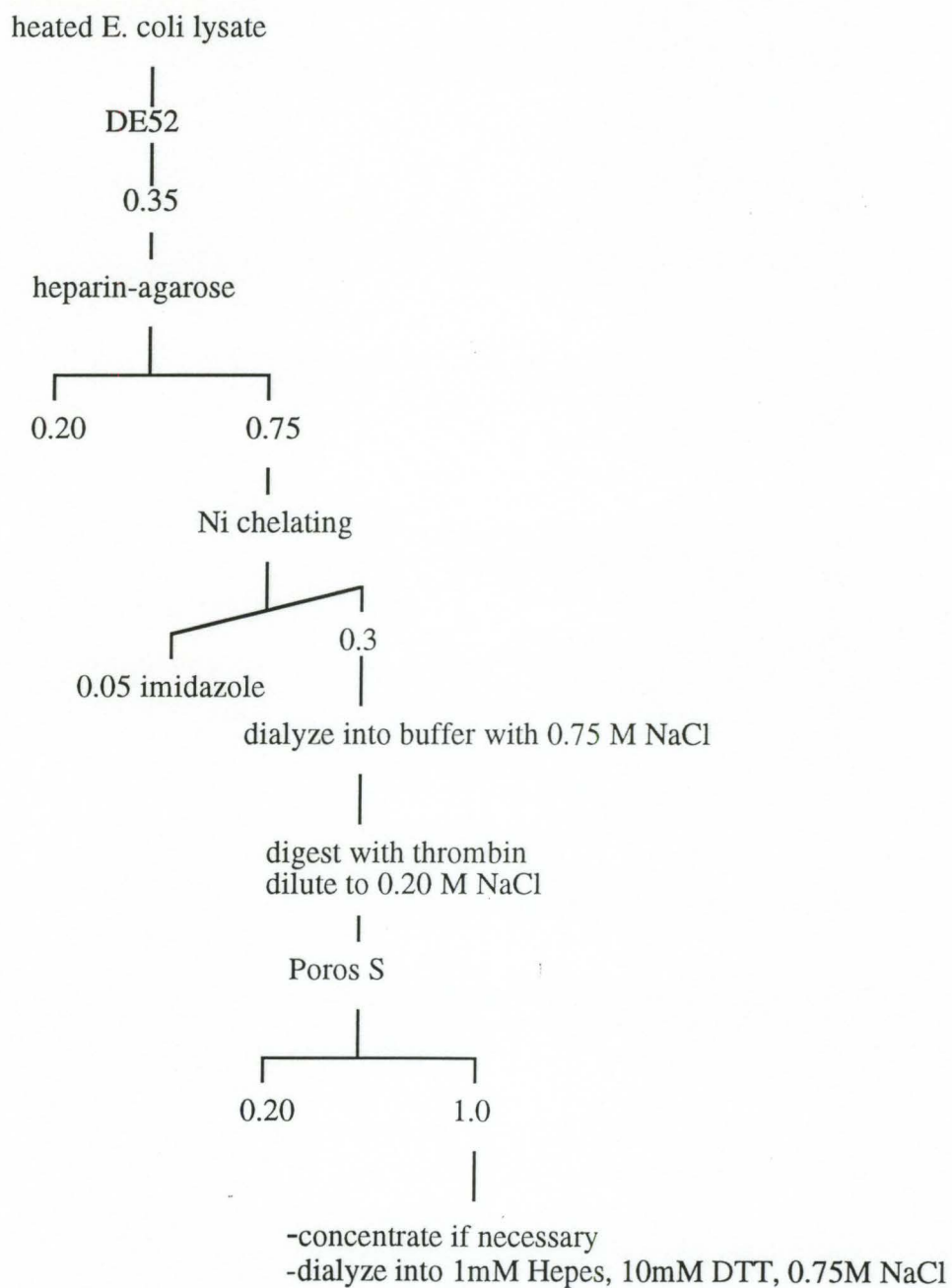


Figure 1b. Purification scheme of EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub>. Horizontal and diagonal lines represent step and gradient elutions, respectively. Elution was with sodium chloride with concentrations given in moles/litre.



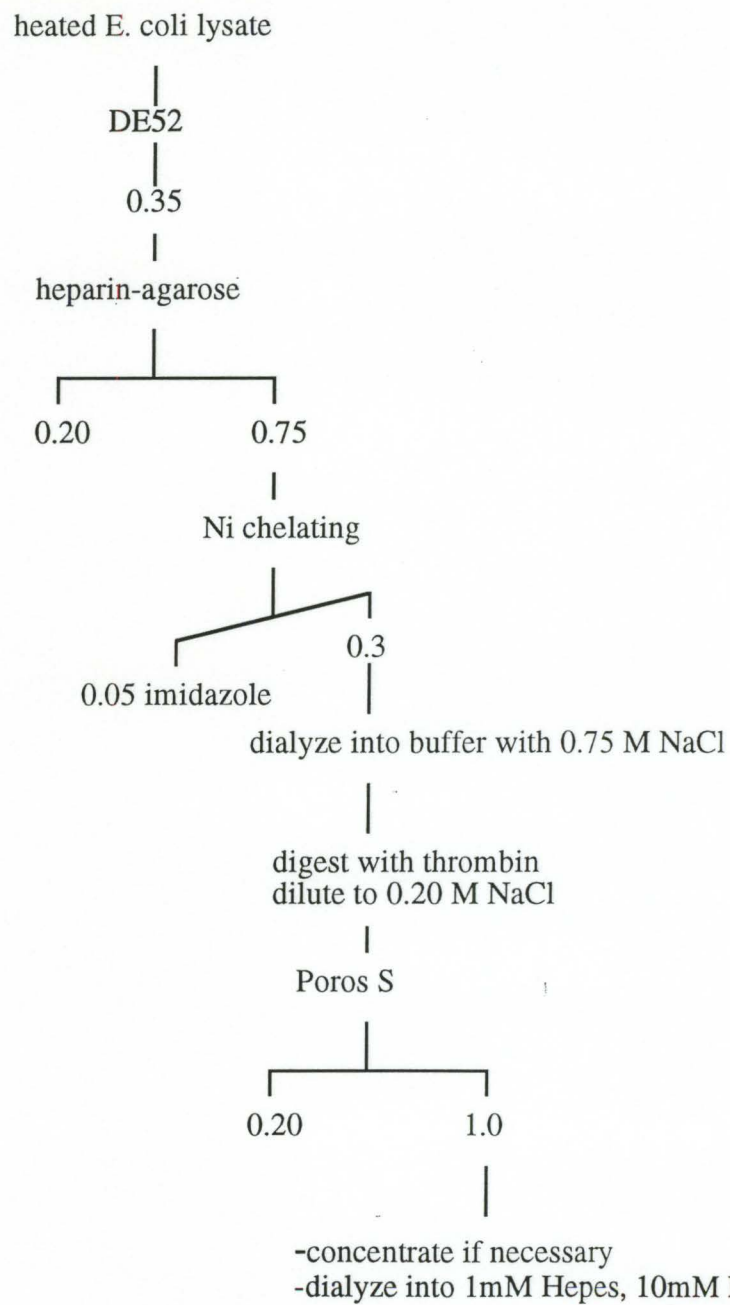


Figure 1c. Purification scheme of EBNA<sub>459-607</sub>. Horizontal and diagonal lines represent step and gradient elutions, respectively. Elution was with sodium chloride with concentrations given in moles/litre.

#### **II.4. Proteolysis of the EBNA-1 truncation mutants**

EBNA<sub>459-619</sub> was subjected to limited protease digestion with trypsin in order to find the smallest active protease resistant piece of EBNA-1. A trial trypsin digestion was done at 37°C with 1 mg of trypsin for every 5, 10 and 25 mg of EBNA<sub>459-619</sub>. Samples were taken at 5, 10, and 25 minutes and the protease reaction stopped by the addition of 0.2 mM PMSF. Products of the digestions were then analyzed by Coomassie staining on a 15% SDS polyacrylamide gel.

#### **II.5. Electrophoretic Mobility Shift Assays**

DNA-protein binding reactions were performed at room temperature. Each incubation mixture contained 50 mM Hepes pH 7.5, 5 mM MgCl<sub>2</sub>, 2 mg sheared herring sperm DNA, 10 fmol radioactive probe, water, NaCl to a final concentration of 200 mM, and protein to a final volume of 20 ml. Protein extract was added last and the mixtures were incubated for 10 minutes before adding 10 % glycerol and loading directly onto a 4 % polyacrylamide gel in 1 x TBE (90 mM Tris, 64.6 mM boric acid, 2.5 mM EDTA, pH 8.0). After electrophoresis, gels were visualized by autoradiography on Kodak XAR-5 film.

## II.6. Preparation of Radioactive Dyad Symmetry element

The plasmid pGEM-dyad which contains the dyad symmetry element, was digested with BamHI and EcoRI (New England Biolabs) and the 317 base pair insert was purified by the GeneClean method (Bio101 Inc.). To radio-label this fragment, 2 pmol of the DNA was mixed with 0.075 mM of each of dCTP, dGTP, dTTP, 30 mCi of a  $^{32}\text{P}$  ATP, 10 mM Tris-HCl pH 7.5, 5 mM  $\text{MgCl}_2$ , 7.5 mM dithiothreitol and 1 unit of Klenow (New England Biolabs) to a final volume of 20  $\mu\text{l}$ . This mixture was incubated for 10 minutes at  $37^\circ\text{C}$  and purified by G-25 size exclusion chromatography (Pharmacia).

## II.7. Crystallization

I searched for crystallization conditions for each EBNA-1 truncation mutant firstly by the hanging drop method using Crystal Screen I and II (Hampton Research) (Jancariak and Kim, 1991) at room temperature. Once suitable small crystals were found, I expanded these conditions by using different concentrations of the precipitant, different pH's and different temperatures ( $4^\circ\text{C}$ ,  $15^\circ\text{C}$  and room temperature).

## II.8. Data Collection

X-ray diffraction was measured at room temperature on an RAXIS II image plate area detector with  $\text{CuK}\alpha$  radiation from a Rigaku RU200 rotating anode generator. The X-rays were focussed to a 0.2 mm spot with Supper double focussing mirrors.



### **III. RESULTS**

#### **III.1. Expression of the EBNA-1 truncation mutants**

Fragments of the EBNA-1 gene were amplified from the plasmid p205 and inserted into the pET15b expression vector. The restriction sites NdeI and BamHI were added to the 5' and 3' ends of the PCR product respectively and used to insert the EBNA-1 fragments between the NdeI and BamHI sites of the T7 polymerase expression vector, pET15b (Novagen). The proteins were thus expressed as C-terminal fusions to an N-terminal 6 histidine tag and a thrombin protease site. Digestion of each fusion protein with thrombin resulted in the inclusion of four residues, glycine, serine, histidine, methionine, at the N-terminus of the EBNA1 derivatives (see Figure 2).

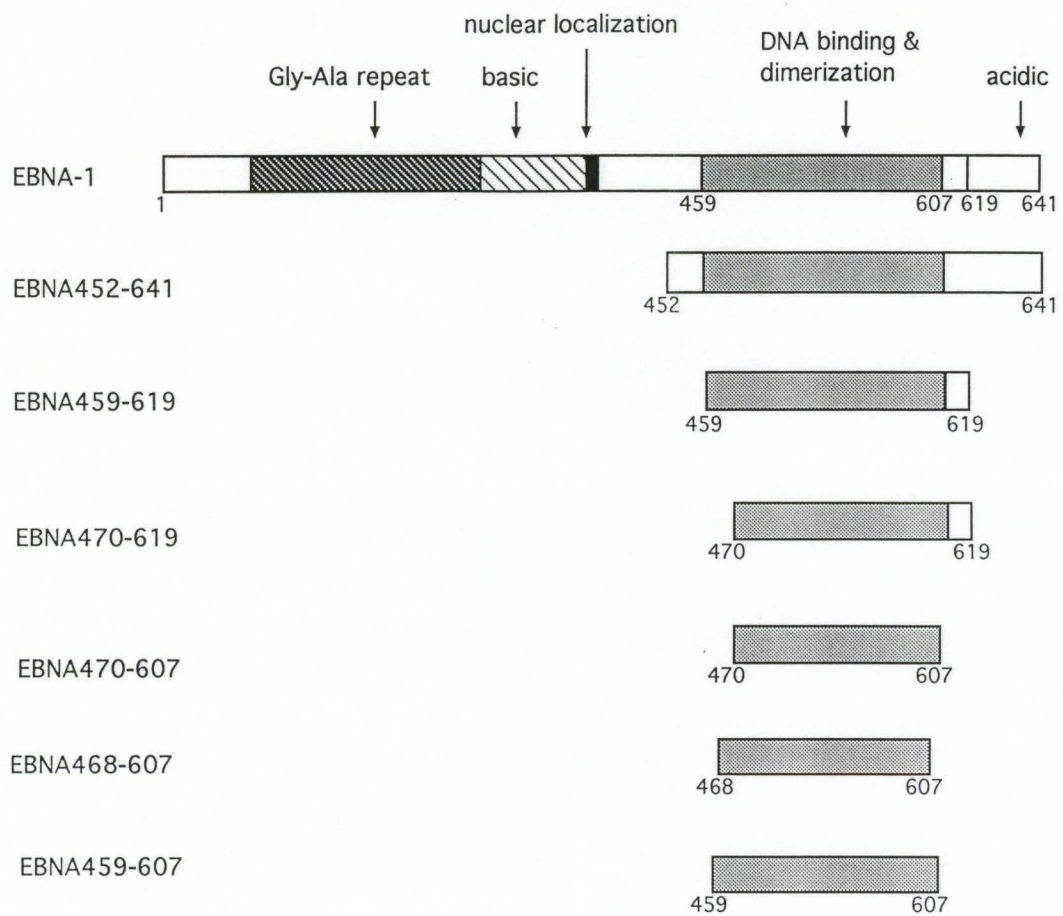


Figure 2. EBNA-1 truncation mutants. EBNA-1 amino acids present in each truncation mutant and some of the salient features of EBNA-1 are shown.

### III.2. Purification

The purification of the six truncation mutants of EBNA-1 from 30 L of *E. coli* cells is summarized in Table 1. Cells were harvested 3 hours post-induction and the clarified lysate was heated in a flask in a water bath. The lysate was brought to the target temperature and put on ice once protein aggregation was observed. EBNA<sub>452-641</sub> and EBNA<sub>459-619</sub> were loaded on the heparin-agarose column in 350 mM NaCl. The other mutants were loaded on the heparin-agarose column in 200 mM NaCl due to the fact that they flowed through the column at 350 mM NaCl. The EBNA-1 mutants were all eluted from the heparin column with 750 mM NaCl. At this point, the protein appeared to be 20% pure. The eluate was then dialyzed into buffer containing no DTT and loaded on the metal-chelating column charged with nickel. The nickel chelating column was eluted using 300 mM imidazole to yield 90-100% pure mutant EBNA-1. DTT (10 mM) was added to the protein immediately after it eluted from the column and the protein was kept at room temperature in order to avoid aggregation. The protein was dialyzed at 4°C into buffer with 750 mM NaCl, 10 mM DTT, no imidazole, and no protease inhibitors. The protein was then digested with thrombin. To ensure that the thrombin digest went to completion, 10 units per mg of EBNA-1 was added, the protein was incubated at 37°C, and after 1 hour, another 5 units per mg of EBNA-1 was added. The digest was then continued for another 2 hours. Completion of the digestion was confirmed by



Coomassie staining on a 15 % SDS polyacrylamide gel. If the digest was not complete, another 5 units of thrombin per mg of EBNA-1 was added, and the mixture was incubated for another hour. Once the protein was completely digested, it was concentrated on a Poros S column to approximately 10 mg/ml. The Poros S column was eluted at a slow flow rate (1 ml/min) to ensure a high protein concentration in the eluate. If necessary, the protein was concentrated on a Centricon-10 micro-concentrator (Amicon). This step was avoided if possible because some EBNA-1 binds to the membrane. EBNA<sub>452-641</sub>, EBNA<sub>459-619</sub> and EBNA<sub>470-619</sub> were dialyzed into buffer containing 500 mM NaCl. EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub> were dialyzed into buffer containing 750 mM NaCl, and EBNA<sub>459-607</sub> was dialyzed into buffer containing 1 M NaCl. High salt was necessary to maintain solubility. The purified proteins are shown in Figure 3.

### **III.3. The Search for the Most Suitable Fragment of EBNA-1 for Crystallization**

Proteolysis of EBNA<sub>459-619</sub> was performed in an effort to find a tightly folded, protease resistant fragment of EBNA-1 for crystallization. Two slightly smaller, yet active, trypsin-resistant fragments were found; N-terminal sequencing revealed that the major peptides began at amino acid 470 with a minor component beginning at amino acid 468. Electrospray mass spectroscopy



indicated that the mass difference between EBNA<sub>459-619</sub> and the proteolytic products were due only to N-terminal truncations (see Figure 4) Resistance to trypsin digestion suggested that these EBNA fragments were more tightly folded and therefore good candidates for crystallization. These studies pointed to amino acid 470 or amino acid 468 as the approximate N-terminal boundary of the DNA binding domain. The C-terminal boundary could not be as well defined in these studies due to the lack of suitable protease sites near amino acid 619.

Table 1  
Purification of EBNA-1 Truncation Mutants

Truncation mutant	Fraction	Protein (approx.) <i>mg</i>	Truncation mutant	Fraction	Protein (approx.) <i>mg</i>
EBNA452-641	lysate	8000	EBNA468-607	lysate	5500
	heated lysate	3500		heated lysate	2300
	DE52	2800		DE52	1350
	heparin-agarose	550		heparin-agarose	50
	nickel-chelating	120		nickel-chelating	35
	Poros S	100		Poros S	24
EBNA459-619	lysate	8000	EBNA470-607	lysate	6500
	heated lysate	4500		heated lysate	3500
	DE52	2500		DE52	1800
	heparin-agarose	500		heparin-agarose	150
	nickel-chelating	60		nickel-chelating	14.1
	Poros S	50		Poros S	9
EBNA470-619	lysate	7500	EBNA459-607	lysate	1823
	heated lysate	4900		heated lysate	1412
	DE52	2800		DE52	760
	heparin-agarose	450		heparin-agarose	120
	nickel-chelating	191		nickel-chelating	10.5
	Poros S	140		Poros S	8

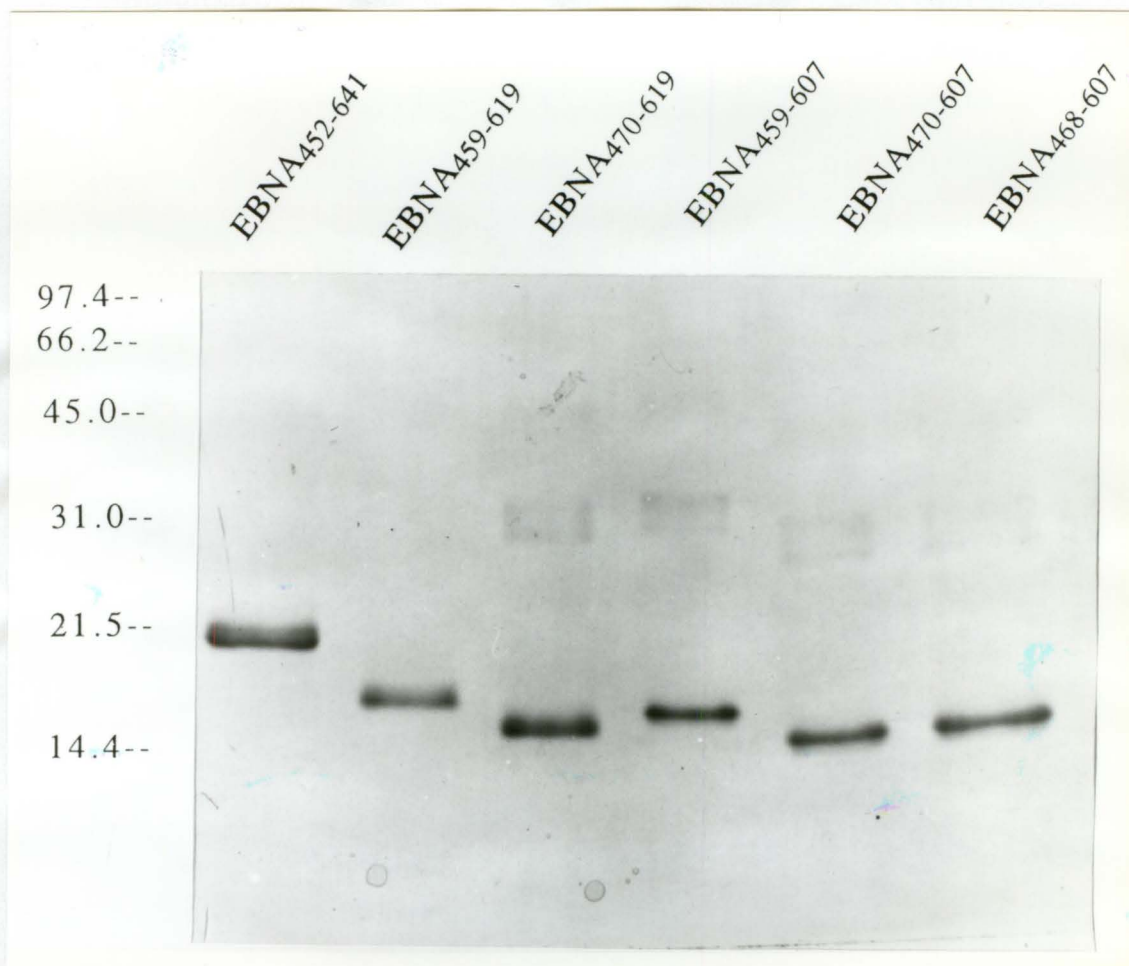


Figure 3. SDS polyacrylamide gel of the EBNA-1 truncation mutants. 1.0  $\mu\text{g}$  of protein purified from *E. coli* lysates was subjected to electrophoresis on a 15% polyacrylamide gel containing SDS and stained with Coomassie blue. Higher molecular weight species seen in some lanes are EBNA-1 dimers, not impurities. Molecular weights in kilodaltons are shown to the left of the gel.

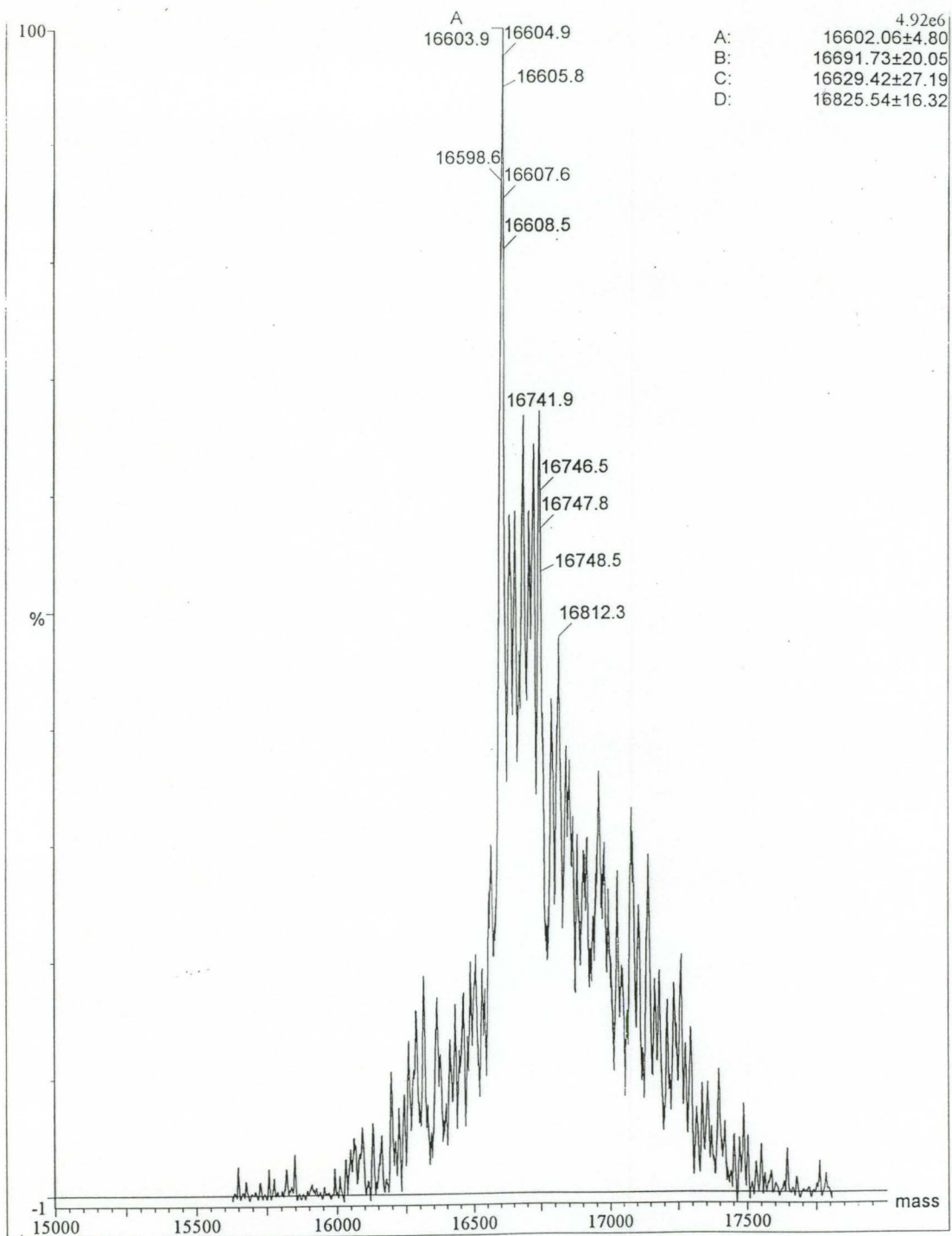


Figure 4. Electrospray mass spectroscopy spectrum for EBNA<sub>470-607</sub>. Peaks are labelled with molecular weights in Daltons.



The C-terminal boundary of the DNA binding domain was selected on the basis of previous studies that indicated that the DNA binding/dimerization region extended only to amino acid 607 (Ambinder et al.,1991). We therefore generated and purified the truncation mutants EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub>.

All of the EBNA-1 fragments were shown by electrophoretic mobility shift assays (see Figure 5) and methylation protection footprints to bind specifically to the 18 bp EBNA-1 recognition sequence (Summers et al., *in preparation*).

#### **III.4. Protein Crystallization**

EBNA<sub>452-641</sub>, EBNA<sub>459-619</sub>, EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub> were crystallized using the method of hanging-drop vapour-diffusion using Linbro multiwell tissue culture plates. The crystal forms were grown at room temperature from an initial protein solution of 10 mg/ml and the hanging drop was comprised of an equal volume of the protein solution and the solution containing the precipitant. All solutions, including the protein, were filtered through 0.22  $\mu$ M filters before crystallization.

To crystallize EBNA<sub>452-641</sub>, the protein was mixed with an equal volume of 50 mM MES pH 6.0, 0.1 M ammonium phosphate and incubated over a reservoir containing the same solution. Hexagonal plates with dimensions 0.1 mm x 0.1 mm x 0.02 mm (see Figure 6a),

or rectangular rods with dimensions 0.05 mm x 0.05 mm x 0.01 mm (see Figure 6b) grew in 1-7 days. Smaller hexagonal plates could also be induced to grow from 0.45 mg/ml 1-stearyl-2-oleoyl-phosphatidyl choline and 0.05 mg/ml sphingosine lipid layers on the cover slip (see Figure 6c) (Hemming et al., 1995). We were unable to generate EBNA<sub>452-641</sub> crystals that were large enough for X-ray diffraction.

Several crystal forms of EBNA<sub>470-619</sub> were grown. Hexagonal crystals (.2 x .2 x .2 mm) grew at room temperature in 4-6 weeks over a reservoir containing 0.5 ml of a solution containing 50 mM MES pH 6.5, 0.7-0.9 M NaAc (see Figure 6d). Another crystal form (0.25 x 0.05 mm x 0.05 mm rectangular rods) grew over a reservoir of 50 mM HEPES pH 7.0-7.5, 7-10% PEG 4K, 500 mM NaCl (see Figure 6e). Diffraction analysis revealed that these crystal forms were twinned and not ideal for X-ray analysis.

Two other crystal forms of EBNA<sub>470-619</sub> grew in 4-6 weeks over a reservoir solution containing 50 mM MES pH 6-6.5 and 0.8-1.0 M sodium formate (see Figure 6f). Both crystals had a hexagonal morphology and grew to a maximal dimension of 0.2 x 0.2 x 0.2 mm. The diffraction from each of these crystals extended to beyond 2.5 Å resolution. These two crystal forms could only be distinguished by diffraction analysis. The first crystal was trigonal with a unit cell of  $a=b=86.7\text{Å}$   $c=31.8\text{Å}$  and the space group assigned as P312. Only two crystals with this space group were detected and, in these instances, co-existed in the same drop with another crystal form, which was

also trigonal with a very similar unit cell of dimensions  $a=b=86.5\text{\AA}$   $c=31.8\text{\AA}$ . The space group of the second crystal form was tentatively assigned as P3. A heavy atom derivative of the P3 crystals has been prepared by soaking the crystals in Hg-acetate, and the space group assignment was confirmed.

Crystals of the EBNA<sub>470-607</sub> and EBNA<sub>468-607</sub> were grown over a reservoir solution containing 50 mM MES pH 6.0 and 0-100 mM NaCl. Rectangular rods with dimensions 0.5-1.0 mm x 0.2 mm x 0.2 mm grew at room temperature in 2-3 days (see Figure 6g). X-ray diffraction analysis showed that these crystals were tetragonal with unit cell dimensions  $a=59\text{\AA}$ ,  $b=66\text{\AA}$  and  $c=69\text{\AA}$ ; the space group assignment was  $P2_12_12_1$ . The diffraction extended to beyond  $2\text{\AA}$  and a heavy atom derivative was obtained by soaking in a solution containing pCMB.

Native data sets were collected from the  $P2_12_12_1$  crystals both at room temperature and at  $-175^\circ\text{C}$ . The low temperature was maintained using a nitrogen gas delivery system from Molecular Structure Corporation. The crystals were cryoprotected by sequential 5 minute incubations of the crystal in the mother liquor appended with 5, 10, 15, and 20% glycerol, mounted in a loop made from a strand of dental floss and then flash frozen in the stream of nitrogen gas.



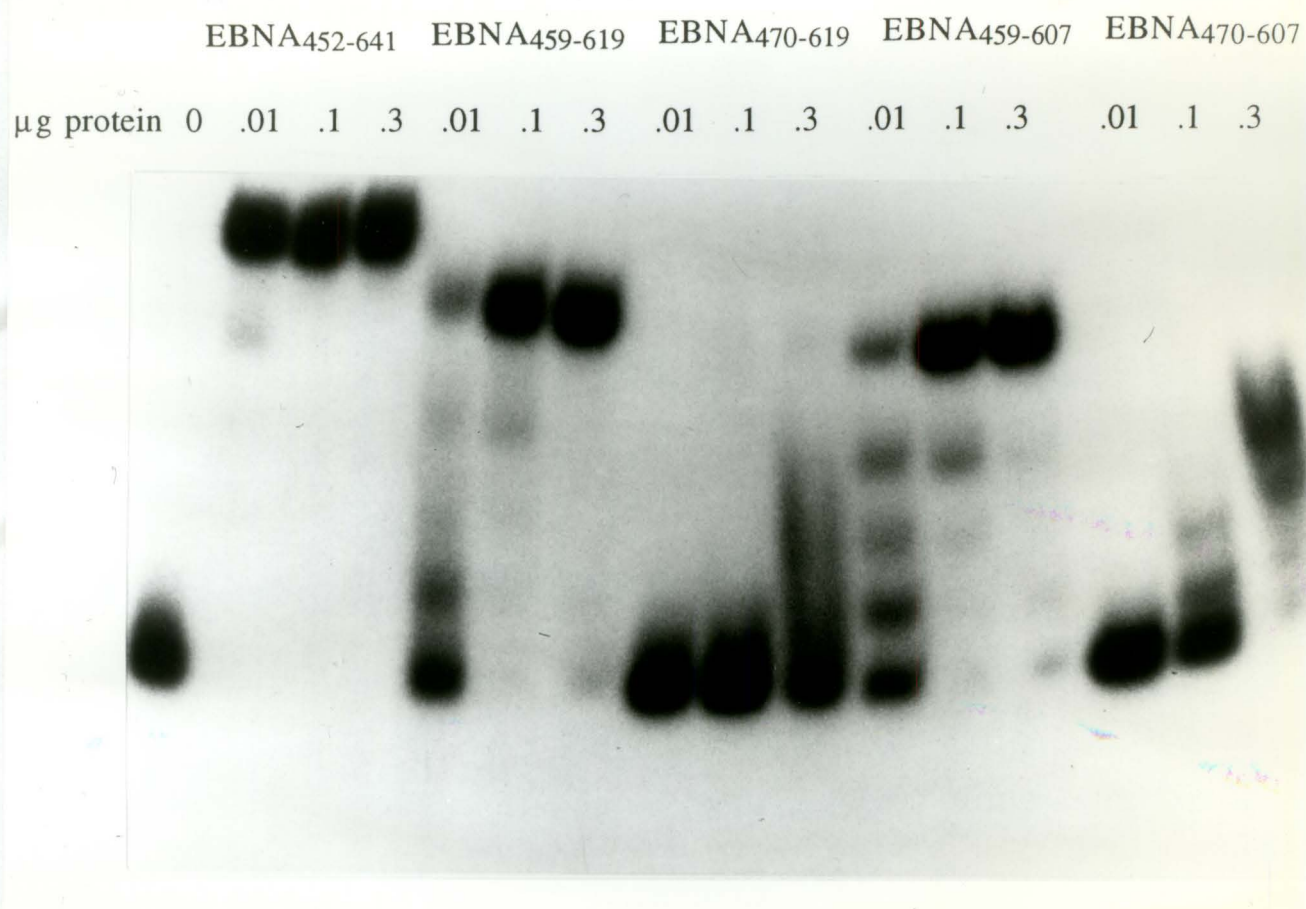


Figure 5. Mobility shift assay using 10 fmol of a 317 base pair insert from the plasmid pGEM-dyad as a probe and indicated amount of the bacterially expressed truncation mutants of EBNA-1. The probe contains the sequence corresponding to the Dyad symmetry element of *oriP*. The first lane shows the position of the probe DNA in the absence of EBNA-1.



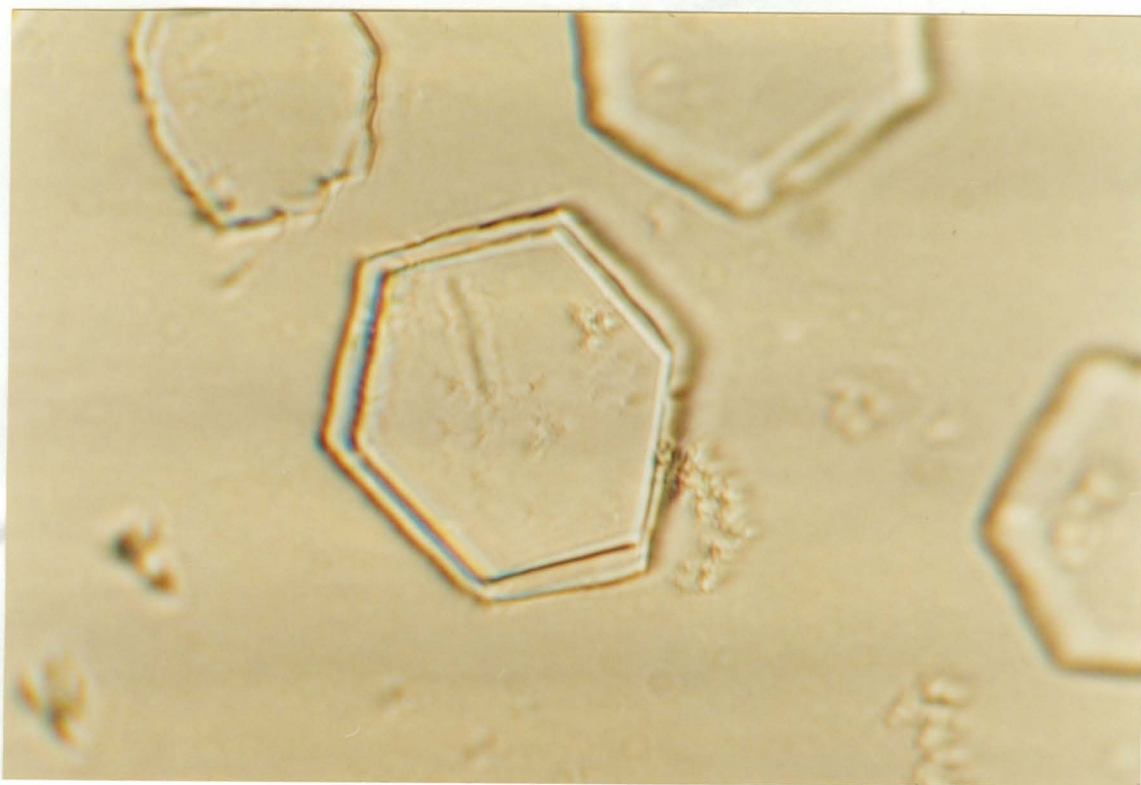


Figure 6a. Crystals of EBNA<sub>452-641</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown at room temperature over a reservoir solutions containing 50 mM MES pH 6.0, 0.1 M ammonium phosphate to the dimensions 0.1 mm x 0.1 mm x 0.02 mm in 1-7 days.

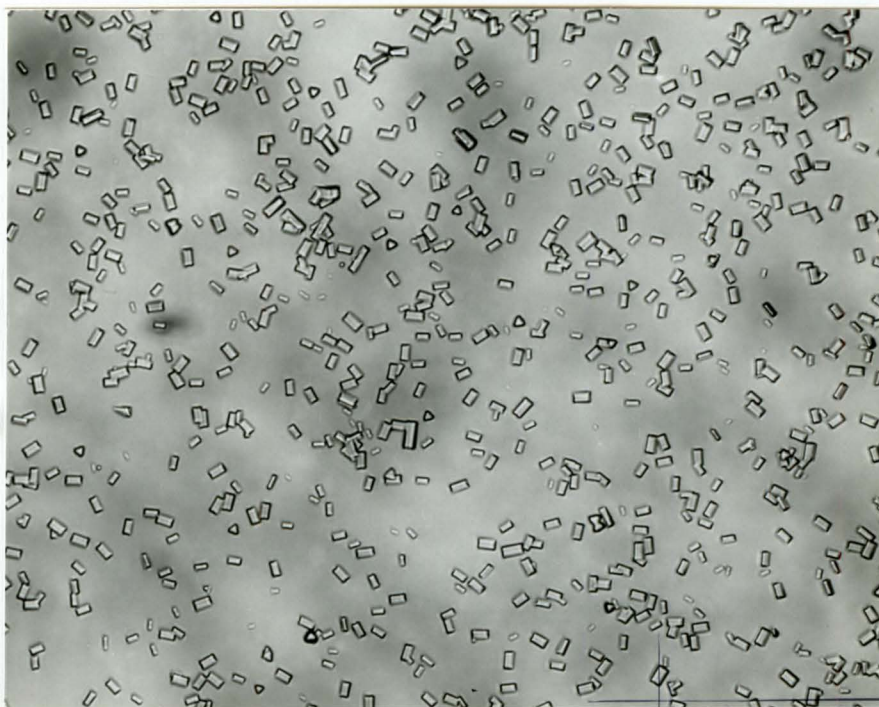


Figure 6b. Crystals of EBNA<sub>452-641</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown under the same conditions as the hexagonal plates shown in figure 6a. Dimensions are 0.05 mm x 0.05 mm x 0.01 mm.

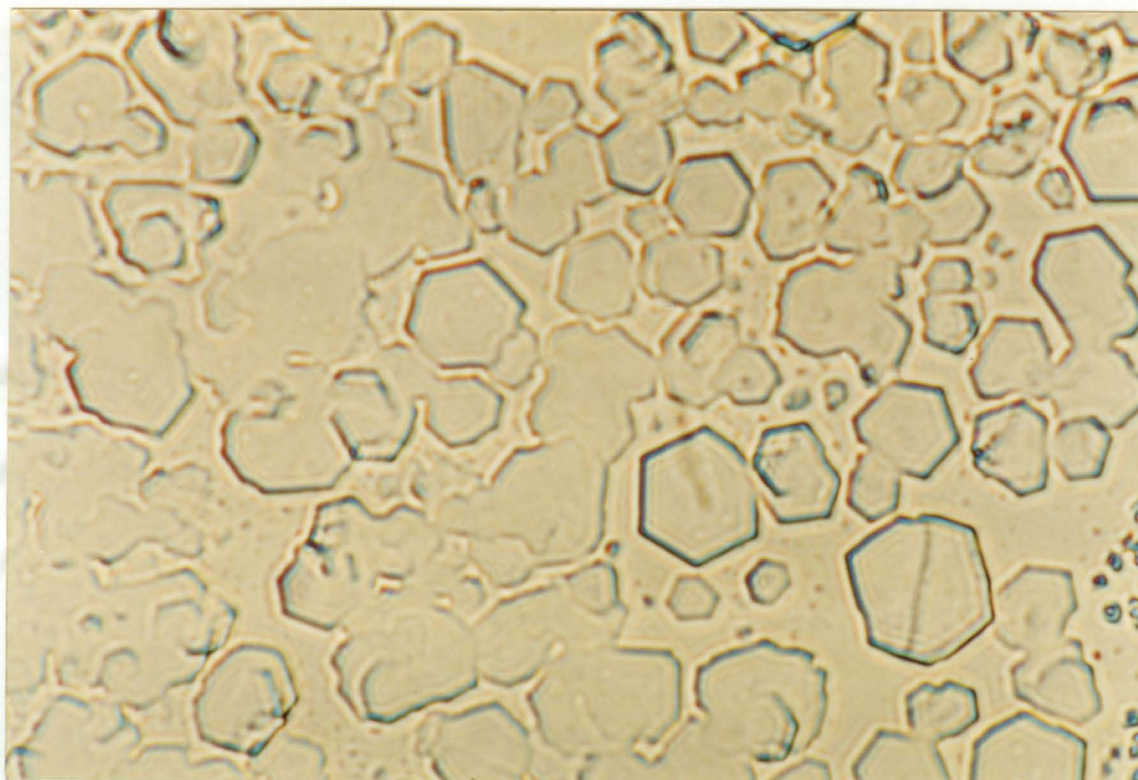


Figure 6c. Crystals of EBNA<sub>452-641</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown under the same conditions as the hexagonal plates shown in figure 6a, except that they were grown on 1-stearyl-2-oleoyl-phosphatidylcholine/sphingosine lipid layers on the cover slip.



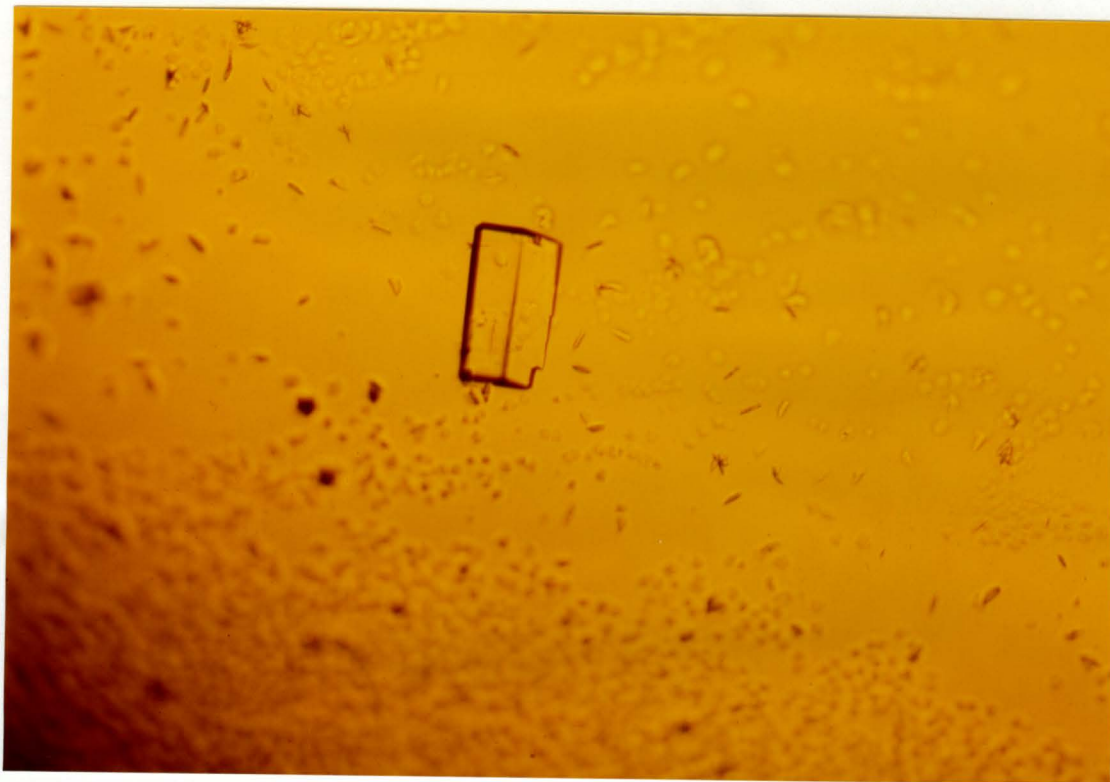


Figure 6d. Crystals of EBNA<sub>470-619</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown at room temperature over a reservoir solution containing 50 mM MES pH 6.5, and 0.7-0.9 M NaAc to the dimensions 0.2 x 0.2 x 0.2 mm.



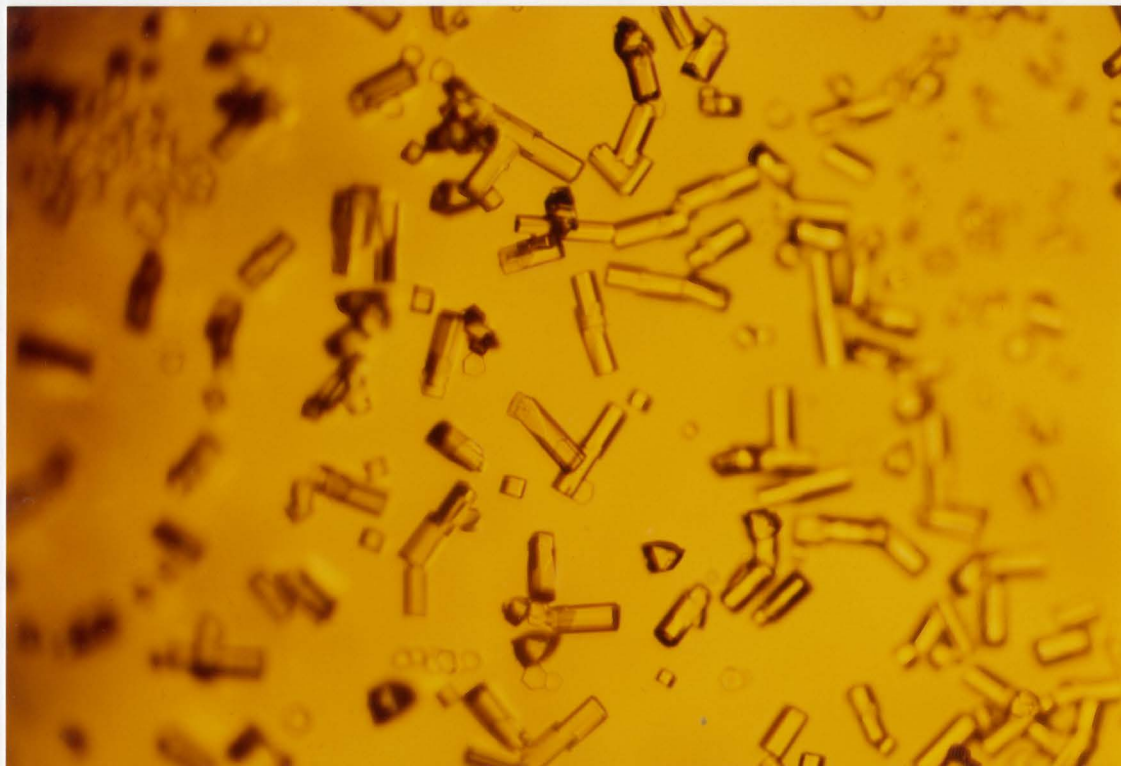


Figure 6e. Crystals of EBNA<sub>470 619</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown at room temperature over a reservoir solution containing 50 mM HEPES pH 7.0-7.5, 7-10% PEG 4K, and 500 mM NaCl to the dimensions 0.25 x 0.05 x 0.05 mm.

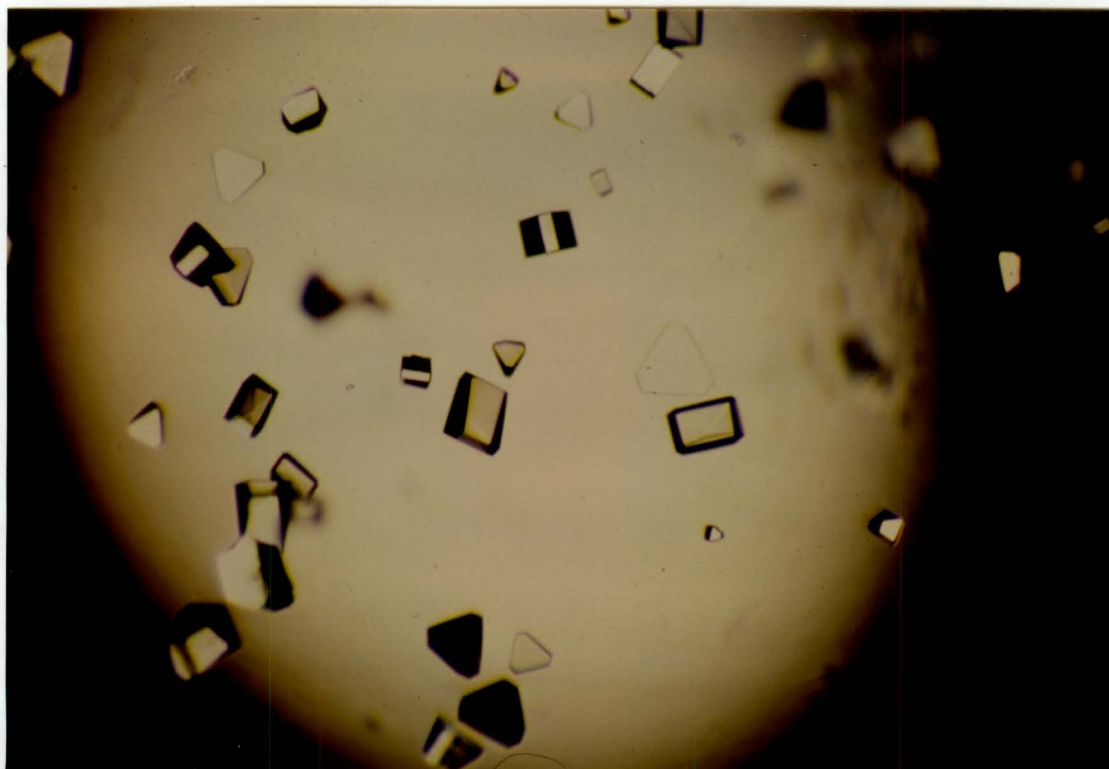


Figure 6f. Crystals of EBNA<sub>470-619</sub> grown by the hanging-drop vapour diffusion method. Crystals were grown at room temperature over a reservoir solution containing 50 mM MES pH 6-6.5, and 0.8-1.0 M sodium formate to the dimensions 0.2 x 0.2 x 0.2 mm.

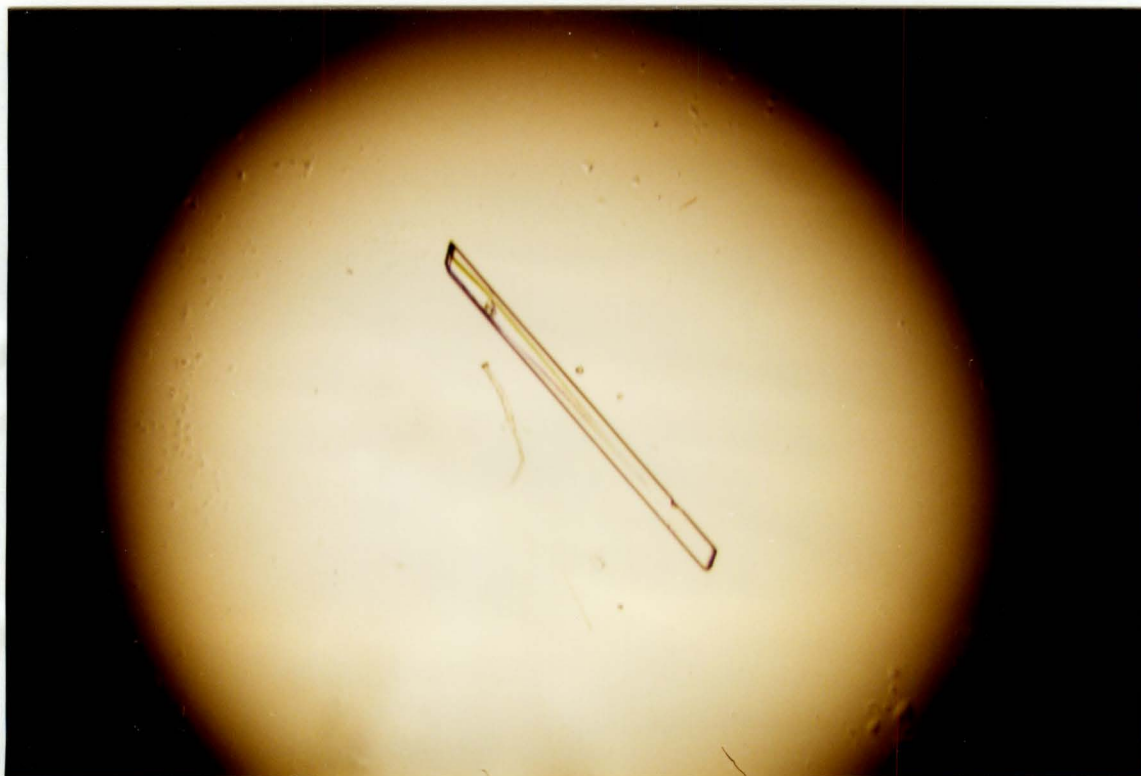


Figure 6g. Crystal of EBNA<sub>470-607</sub> grown by the hanging drop vapour diffusion method. Crystals were grown over a reservoir solution containing 50 mM MES pH 6.0 and 0-100 mM NaCl at room temperature to the dimensions 0.5-1.0 x 0.2 x 0.2 mm.



#### IV. DISCUSSION

The crystallization of a protein is often approached by the rigorous screening of thousands of different crystallization conditions. Variables include different precipitants, pH, temperatures, protein concentration, and additives such as detergents, alcohols, and glycerol. It is a very time-consuming and tedious process with no guarantee of results. For example, many groups have been trying to crystallize the origin binding protein of SV40, T-antigen, for many years and only succeeded in growing small crystals, not suitable for X-ray analysis. In this study, I have reported the first crystals of an origin binding protein which are suitable to be studied by X-ray crystallography. This accomplishment is not only biologically significant, but also significant from a technical point of view because of the approach that was taken to find the most suitable truncation mutant for crystallization. I suggest that the search for the best piece of the protein for crystallization is as important as the search for crystallization conditions. Indeed, once the correct protein fragment is found, finding crystallization conditions becomes considerably easier.

It has been understood for many years that proteins are composed of domains. The domain is "part of the polypeptide chain which can independently fold into a stable tertiary structure"



(Brandon and Tooze, 1991). The formation of a well-ordered crystal depends on the ability of the protein to pack efficiently and tightly. Extraneous pieces of the polypeptide that do not fold into the domain hinder the ability of the protein to pack into the crystal. Therefore, the crystallographer must try to find almost the exact amino acids where a domain begins and ends. This effort will greatly increase the chance of forming high quality crystals.

I have used limited proteolytic digestion to re-define the borders of the DNA binding/dimerization domain of EBNA-1. Resistance to a protease indicates that a polypeptide is folded very tightly thereby protecting the peptide from digestion. After attempting to crystallize EBNA<sub>452-641</sub>, and obtaining only small, poor quality crystals, I decided to express and purify a smaller piece, EBNA<sub>459-619</sub>. A similar mutant EBNA<sub>459-617</sub>, had been found to be resistant to protease K digestion when bound to DNA (Shah et al., 1992). After purifying EBNA<sub>459-619</sub>, I exposed this protein to limited trypsin digestion and found a slightly smaller, trypsin-resistant piece.

The next step was to determine which amino acids composed this protease resistant piece of the protein. N-terminal sequencing revealed that the polypeptide began at amino acid 470 and electrospray mass spectroscopy confirmed that the mass difference between this polypeptide and EBNA<sub>459-619</sub> was due only to this truncation. Therefore, EBNA<sub>470-619</sub> was cloned and purified for crystallization trials. High quality crystals were obtained and

diffraction data was collected, however this protein packed into crystals of the space group P3. No other protein has been solved in this space group by isomorphous replacement and our data proved too difficult to analyze.

I discovered that even when high quality crystals are obtained, one should not abandon the search for new conditions. It is ideal to have several crystal forms from which to choose, in the event that one crystal form proves unsuitable or that heavy atom derivatives cannot be found with one crystal form. A study by Shah et al. (1992) provided biochemical evidence that the DNA binding/dimerization domain only extended at the C-terminus to amino acid 607. Therefore, EBNA<sub>470-607</sub> was cloned, purified and rigorously tested for suitable crystallization conditions. This truncation mutant was the easiest to crystallize. High quality crystals in the space group  $p2_12_12_1$  were obtained in only 1-2 days simply by lowering the NaCl concentration by vapour diffusion. The ease in which this protein crystallized lends strength to my contention that finding the ideal piece of the protein greatly facilitates the search for crystallization conditions.

Lessons learned from the crystallization of EBNA-1 should help not only in the crystallization of new proteins, but in the crystallization of EBNA-1 when bound to DNA. This project poses a new problem because many DNA binding proteins change shape when bound to DNA (reviewed in Brandon and Tooze, 1991). Moreover, it is important to choose a DNA fragment of the



appropriate length for proper packing into the crystal. Currently, crystal trials with all of the EBNA-1 truncation mutants and eight different oligonucleotides are being pursued.

Biologically, the crystallization of the first origin binding protein has tremendous significance. Very little is known about how origin recognition proteins work, how they attract replication machinery or how they may prepare the origin for replication. EBNA-1 is an excellent model system for the study of mammalian origins of replication and origin binding proteins, because the latent EBV genome is replicated once every cellular S phase (Adams, 1987; Yates and Guan, 1991). Knowledge of how EBNA-1 interacts with the cellular replication machinery may lend some understanding to the mechanism of action of mammalian origins and replication proteins.

A putative origin binding protein in eukaryotic cells has been identified and it seems to have some similarity to characterized OBPs, including EBNA-1. The origin recognition complex (ORC) was purified from *S. cerevisiae* cells and contains six different polypeptides. Like  $\lambda$  O protein and EBNA-1, ORC induces DNA distortion, indicated by a periodic pattern of DNaseI hyper and hypo-sensitive sites. This suggests that the DNA may be wrapped around ORC. Like T-antigen from SV40, ORC binds to DNA in an ATP dependent manner. Also, studies have shown that ORC is probably bound to the DNA throughout the cell cycle, like EBNA-1, suggesting that ORC does not regulate DNA replication, but acts to attract the replication machinery. The same role has been deduced for EBNA-1.

Extensive studies of simple systems, like EBV and EBNA-1, will lend great understanding to how DNA replication occurs in eukaryotic cells. Knowledge of the structure of origin binding proteins is crucial for discovering how these systems work. Along with providing information about the approach one should take when crystallizing a new protein, this study will provide the groundwork for understanding how EBNA-1 acts to initiate DNA replication.



**REFERENCES**

- Adams, A. 1987. Replication of latent Epstein Barr virus genomes in Raji cells. *J. Virol.* **61**: 1743-1746.
- Ambinder, R.F., Mullen, M., Chang, Y., Hayward, G.S. and Hayward, S.D. 1991. Functional domains of Epstein-Barr virus nuclear antigen EBNA-1. *J. Virol.* **65**: 1466-1478.
- Borowiec, J.A., Dean, F.B., Bullock, P.A., and Hurwitz, J. 1990. Binding and unwinding-how T antigen engages the SV40 origin of DNA replication. *Cell* **60**: 181-184.
- Brandon, C. and Tooze, J. 1991. Introduction to protein structure. Garland Publishing, Inc. New York and London.
- Bell, S.P. and Stillman, B. 1992. ATP dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**: 128-134.

Chen, M., Middeldorp, J.M., and Hayward, S.D. 1993. Separation of the complex DNA binding domain of EBNA-1 into DNA recognition and dimerization subdomains of novel structure. *J. Virol.* **67**: 4875-4885.

DePamphilis, M.L. 1993. Eukaryotic DNA replication: anatomy of an origin. *Annu. Rev. Biochem.* **62**: 29-63.

Frappier, L. and O'Donnell, M. 1991a. Overproduction, purification, and characterization of EBNA-1, the origin binding protein of EBV. *J. Biol. Chem.* **266**: 7819-7826.

Frappier, L., and O'Donnell, M. 1991b. Epstein-Barr virus nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein-Barr virus. *Proc. Natl. Acad. Sci. USA* **88**: 10875-10879.

Frappier, L., and O'Donnell, M. 1992. EBNA-1 distorts oriP, the Epstein-Barr virus latent replication origin. *J. Virol.* **66**: 1786-1790.

Frappier, L., Goldsmith, K., and Bendell, L. 1994. Stabilization of the EBNA-1 protein on the Epstein-Barr virus latent origin of DNA replication by a DNA looping mechanism. *J. Biol. Chem.* **269**: 1057-1062.

Gahn, T.A., and Schildkraut, C.L. 1989. The Epstein-Barr virus origin of plasmid replication, ori-P, contains both the initiation and termination sites of DNA replication. *Cell* **58**: 527-535.

Gutierrez, C., Guo, Z.S., Roberts, J.M., DePamphilis, M.L. 1990. Simian virus 40 origin auxiliary sequences weakly facilitate T-Antigen binding but strongly facilitate DNA unwinding. *Mol Cell Biol.* **10**: 1719-1728.

Harrison, S., Fisenne, K., and Hearing, J. 1994. Sequence requirements of the Epstein-Barr virus latent origin of DNA replication. *J. Virol.* **68**: 1913-1925.

Hearing, J., Mulhaupt, Y., and Harper, S. 1992. Interaction of Epstein-Barr virus nuclear antigen 1 with the viral latent origin of replication. *J. Virol.* **66**: 694-705.

Hemming, S.A., Bochkarev, A., Darst, S.A., Kornberg, R.D., Ala, P., Yang, D.S.C., and Edwards, A.M. 1995. The mechanism of protein crystal growth from lipid layers. *J. Mol. Biol.* **246**: 308-316.

Hsieh, D-J. Camiolo, S.M. and Yates, J.L. 1993. Constitutive binding of EBNA-1 protein to the Epstein-Barr virus origin, *oriP*, with distortion of DNA structure during latent infection. *EMBO.* **12**: 4933-4944.

Jancariak, J. and Kim, S.H. 1991. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **24**: 409-411.

Jankelevich, S., Kolman, J.L., Bodnar, J.W., and Miller, G. 1992. A nuclear matrix attachment region organizes the Epstein-Barr viral plasmid in Raji cells into a single DNA domain. *EMBO*. **11**: 1165-1176.

Koff, A., Schwedes, J.F., Tegtmeyer, P. 1991. Herpes simplex virus origin-binding protein (UL9) loops and distorts the viral replication origin. *J.Virol.* **65**: 3284-3292.

Kornberg, A. 1992. DNA Replication. Second Edition. W.H. Freeman and Company, N.Y.

Oh, S-J., Chittenden, T., and Levine, A.J. 1991. Identification of cellular factors that bind specifically to the Epstein-Barr virus origin of DNA replication. *J.Virol.* **65**: 514-519.

Polvino-Bodnar, M., Kiso, J., and Schaffer, P.A. 1988. Mutational analysis of Epstein-Barr virus nuclear antigen-1 (EBNA-1). *Nuc. Acids Res.* **16**: 3417-3435.

Polvino-Bodnar, M. and Schaffer, P.A. 1992. DNA binding activity is required for EBNA-1 dependent transcriptional activation and DNA replication. *Virology.* **187**: 591-603.



Shaw, J.E., Levinger, L.F., and Carter, C.W. Jr. 1979 Nucleosomal structure of Epstein-Barr virus DNA in transformed cell lines. *J. Virol.* **29**: 657-665.

Stillman, B. 1994. Initiation of chromosomal DNA replication in eukaryotes. *J. Biol. Chem.* **269**: 7047-7050.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. 1990. *Meth. Enzymol.* **185**: 60-89.

Sugden B., and Warren, N. 1989. A promotor of Epstein-Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. *J. Virol.* **63**: 2644-2649.

Su, W., Middleton, T., Sugden, B., Echols., H. 1991. 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**: 10870-10874.

Waga, S., and Stillman, B. 1994. Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. *Nature.* **369**: 207-212.

Yang, L., Li, R., Mohr, I.J., Clark, R. and Botchan, B. 1991. Activation of BPV-1 replication in vitro by the transcription factor E2. *Nature*. **353**: 628-632.

Yates, J.L., Warren, N., Reisman, D. and Sugden, B. 1984. A *cis*-acting element from Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *Proc. Natl. Acad. Sci. USA* **81**: 3806-3910.

Yates, J.L., Warren, N. and Sugden, B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in a variety of mammalian cells. *Nature* **313**: 812-815.

Yates, J.L. and Camiolo, S.M. 1988. Dissection of DNA replication and enhancer functions of Epstein-Barr virus nuclear antigen 1. *Cancer Cells* **6**: 197-205.

Yates, J.L. and Guan, N. 1991. Epstein Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *J. Virol.* **65**: 483-488.

Zhang, S., and Nonoyama, M. 1994. The cellular proteins that bind specifically to the Epstein-Barr virus origin of plasmid DNA replication belong to a gene family. *Proc. Natl. Acad. Sci. USA* **91**: 2843-2847.