

RESEARCH

Open Access

# Analysis of an ankyrin-like region in Epstein Barr Virus encoded (EBV) BZLF-1 (ZEBRA) protein: implications for interactions with NF- $\kappa$ B and p53

David H Dreyfus<sup>1\*</sup>, Yang Liu<sup>2</sup>, Lucy Y Ghoda<sup>3</sup> and Joseph T Chang<sup>2</sup>

## Abstract

**Background:** The carboxyl terminal of Epstein-Barr virus (EBV) ZEBRA protein (also termed BZLF-1 encoded replication protein Zta or ZEBRA) binds to both NF- $\kappa$ B and p53. The authors have previously suggested that this interaction results from an ankyrin-like region of the ZEBRA protein since ankyrin proteins such as I $\kappa$ B interact with NF- $\kappa$ B and p53 proteins. These interactions may play a role in immunopathology and viral carcinogenesis in B lymphocytes as well as other cell types transiently infected by EBV such as T lymphocytes, macrophages and epithelial cells.

**Methods:** Randomization of the ZEBRA terminal amino acid sequence followed by statistical analysis suggest that the ZEBRA carboxyl terminus is most closely related to ankyrins of the invertebrate cactus I $\kappa$ B-like protein. This observation is consistent with an ancient origin of ZEBRA resulting from a recombination event between an ankyrin regulatory protein and a fos/jun DNA binding factor. *In silico* modeling of the partially solved ZEBRA carboxyl terminus structure using PyMOL software demonstrate that the carboxyl terminus region of ZEBRA can form a polymorphic structure termed ZANK (ZEBRA ANKyrin-like region) similar to two adjacent I $\kappa$ B ankyrin domains.

**Conclusions:** Viral capture of an ankyrin-like domain provides a mechanism for ZEBRA binding to proteins in the NF- $\kappa$ B and p53 transcription factor families, and also provides support for a process termed "Ping-Pong Evolution" in which DNA viruses such as EBV are formed by exchange of information with the host genome. An amino acid polymorphism in the ZANK region is identified in ZEBRA from tumor cell lines including Akata that could alter binding of Akata ZEBRA to the p53 tumor suppressor and other ankyrin binding protein, and a novel model of antagonistic binding interactions between ZANK and the DNA binding regions of ZEBRA is suggested that may be explored in further biochemical and molecular biological models of viral replication.

**Keywords:** p53, NF- $\kappa$ B, transcription, ankyrin, phylogeny, oncogenes, viral carcinogenesis, viral conditioning, epigenomics, ping-pong evolution, ASPP2, ASPP1, iASPP

## Background

Analysis of viruses, mobile DNA sequences and host genomes reveals evidence that genetic information can be shared resulting in generation of novel hybrid genetic elements. For example, the ubiquitous human pathogen Epstein Barr Virus (EBV), genome structure contains a region co-linear with the vertebrate variable

immunoglobulin gene segments including regulatory regions for the myc oncogene, signals for immunoglobulin switch recombination, somatic mutation and coding regions[1,2]. Similarly, the EBV viral termini contain sequences similar to somatic V(D)J immunoglobulin and T cell receptor recombination sites and the virus encodes a recombinase co-regulated with the host RAG V(D)J recombinase [3-5].

These observations suggest that EBV and other DNA viruses evolve through sharing of genetic information with the host genome termed "Ping Pong Evolution"

\* Correspondence: dhfreyfus@pol.net

<sup>1</sup>Department of Pediatrics, Yale SOM, 488 Norton Parkway, New Haven CT 06511, USA

Full list of author information is available at the end of the article

[1,2]. In this work, a remarkable example of Ping Pong Evolution between EBV and the host immune system is characterized. The EBV viral protein ZEBRA, Zta encoded by BZLF-1 is proposed to result from capture and fusion of exons from two different host immune response genes, the *fos/jun* transcription factor and  $I\kappa B$  immune regulatory proteins. Analysis of host  $I\kappa B$  proteins is utilized to present an empirically testable model of ZEBRA binding to NF- $\kappa B$  immune response proteins in which the unstructured carboxyl region of ZEBRA can assume two different structural conformations when bound to NF- $\kappa B$ . In addition, an amino acid polymorphism in this region of ZEBRA is identified that could potentially alter the functional properties of the ZEBRA protein interactions with both NF- $\kappa B$  and the distantly related tumor suppressor p53[6].

Studies of ZEBRA, the lytic switch protein, have demonstrated that the two amino terminal exons of ZEBRA are structurally related to the *fos/jun* transcription factors and that ZEBRA binding to DNA sites, known as ZRE, is sufficient to activate the viral lytic cycle[7,8]. The carboxyl region of ZEBRA protein can bind to components of the NF- $\kappa B$  transcription family *in vitro* and alter NF- $\kappa B$  transcription *in vivo*[9,10]. Inactivation of NF- $\kappa B$  transcription is apparently independent of ZEBRA's ability to function as a transcription factor [11-13]. It is not currently known which of these multiple effects of ZEBRA contribute to viral carcinogenesis in a humanized mouse model [14].

One potential outcome of these interactions is that inactivation of NF- $\kappa B$  transcription during viral lytic replication confers a selective advantage upon the virus [15,16]. NF- $\kappa B$  transcription factors and regulatory  $I\kappa B$  proteins are central mediators of both the innate and acquired immune responses [17,18]. Inactivation of NF- $\kappa B$  could block the innate immune response in B lymphocytes which are the viral host cell. Interaction between ZEBRA and NF- $\kappa B$  can trigger apoptosis of cells expressing ZEBRA [19-21]. Inactivation NF- $\kappa B$  transcription by ZEBRA in host B lymphocytes may also oppose the effects of latency proteins that activate NF- $\kappa B$  transcription, delaying apoptosis and contributing to viral maturation and release from pre-apoptotic host cells.

Other bystander cells such as T-lymphocytes, epithelial cells and macrophages may also be transiently infected with EBV. T lymphocytes in particular express both the EBV CD21 receptor as well as other EBV receptors and express ZEBRA protein [3,21,22]. Studies of the effects of transient and stable expression of ZEBRA in T-lymphoblastoid cell lines have confirmed an  $I\kappa B$ -like inactivation NF- $\kappa B$  signaling by ZEBRA and increased T lymphocyte apoptosis [21]. Thus, an additional selective advantage of interactions between

ZEBRA and NF- $\kappa B$  may be destruction of immune responder cells through transient infection and apoptosis.

Another human gamma herpesvirus, human herpesvirus-8, first identified as the cause of Kaposi's sarcoma, encodes a lytic replication protein which lacks DNA binding but has extensive amino acid similarity to ZEBRA including the carboxyl region of ZEBRA [23-25]. This region of ZEBRA is highly conserved between different viral strains [26]. EBV infected lymphocytes also express a trans-spliced ZEBRA like protein termed RAZ sharing the carboxyl terminal NF- $\kappa B$  binding region of the ZEBRA protein, but lacking the ability to bind DNA or ZRE[15,27]. RAZ is also expressed in EBV infected T lymphocytes[3,22].

Remarkably, the carboxyl region of ZEBRA required for interactions with NF- $\kappa B$  transcription and ZEBRA dimerization also binds to p53 tumor suppressor *in vitro* and alters p53 transcription *in vitro*[28-30]. The effects of ZEBRA on p53 transcription are stimulatory in some cases such as T-lymphocytes and epithelial cells, but inhibitory in B-lymphocytes[29,30]. Interactions between ZEBRA and p53 also involve other p53 binding proteins and are more complex than interactions *in vivo* between ZEBRA and NF- $\kappa B$  [31]. The ability of ZEBRA to interact with both NF- $\kappa B$  and p53 has previously been suggested to result from a cryptic  $I\kappa B$ -like region in the carboxyl terminus of the ZEBRA protein[6,21]. NF- $\kappa B$  and p53 proteins share a common  $I\kappa B$  binding region because they are descendents of a common ancestral transcription factor previously termed "proto p53/NF- $\kappa B$ [6]". Since this hypothesis was proposed, the crystal structure of ZEBRA protein has been partially solved including a portion of the carboxyl terminus of the protein interacting with NF- $\kappa B$  and p53[8,32]. In addition, the interactions between p53 binding ankyrin proteins and NF- $\kappa B$  proteins have been characterized providing additional evidence that regulatory proteins in the ankyrin family, including NF- $\kappa B$  inhibitor proteins related to  $I\kappa B$ , bind to both NF- $\kappa B$  and p53 proteins [33].

The recently available partial structure of the ZEBRA carboxyl terminus, and the structures of  $I\kappa B\alpha$  and apoptosis-stimulating protein of p53 (ASPP2, previously known as p53BP2) ankyrin proteins are analyzed in this work, building on previous similarities noted between these regulatory proteins[6]. First, it is demonstrated that primary amino acid similarities between the alpha helix regions of ZEBRA and various other ankyrin proteins are unlikely to have arisen by chance or independent parallel evolution but instead appear to represent capture and homologous descent of a terminal exon encoding an  $I\kappa B$  domain by ZEBRA. Second, it is shown that the partial crystal structure of ZEBRA is consistent

with dimerization of the terminal region of the protein to form an  $I\kappa B$ -like stem and loop structure. Finally, individual ankyrin domains of  $I\kappa B\alpha$  are shown to have structural similarities to the ankyrins of ASPP2, and thus able to bind to the same region targeted by ZEBRA [6].

## Methods

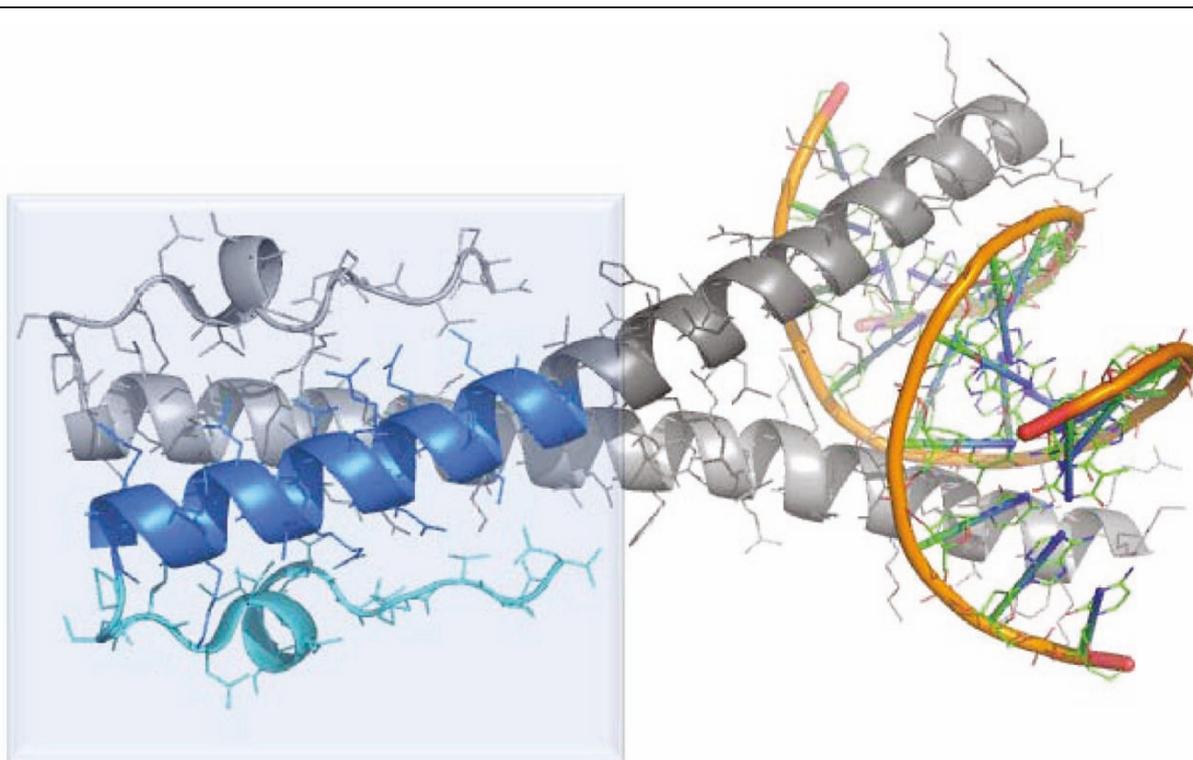
### Definition of the ankyrin-like region of ZEBRA protein termed "ZANK"

We previously noted that a region of ZEBRA (Figures 1, 2) has primary amino sequence similarity to  $I\kappa B$  prior to the determination of the crystal structure of ZEBRA [6]. We first used the program PyMOL (Delano Scientific Inc., CA) to define and visualize a region of ZEBRA whose primary amino acid sequence is similar to ankyrin proteins [6]. ZEBRA structure was partially solved by x-ray crystallography in 2006 and structural coordinates are available as a protein data-base in the public domain (2C9N.pdb). This region of ZEBRA, termed in this work ZANK "ZEBRA ANKryn-like region" has an alpha-helical stem and carboxyl unstructured region (Figure 1). In the course of this analysis,

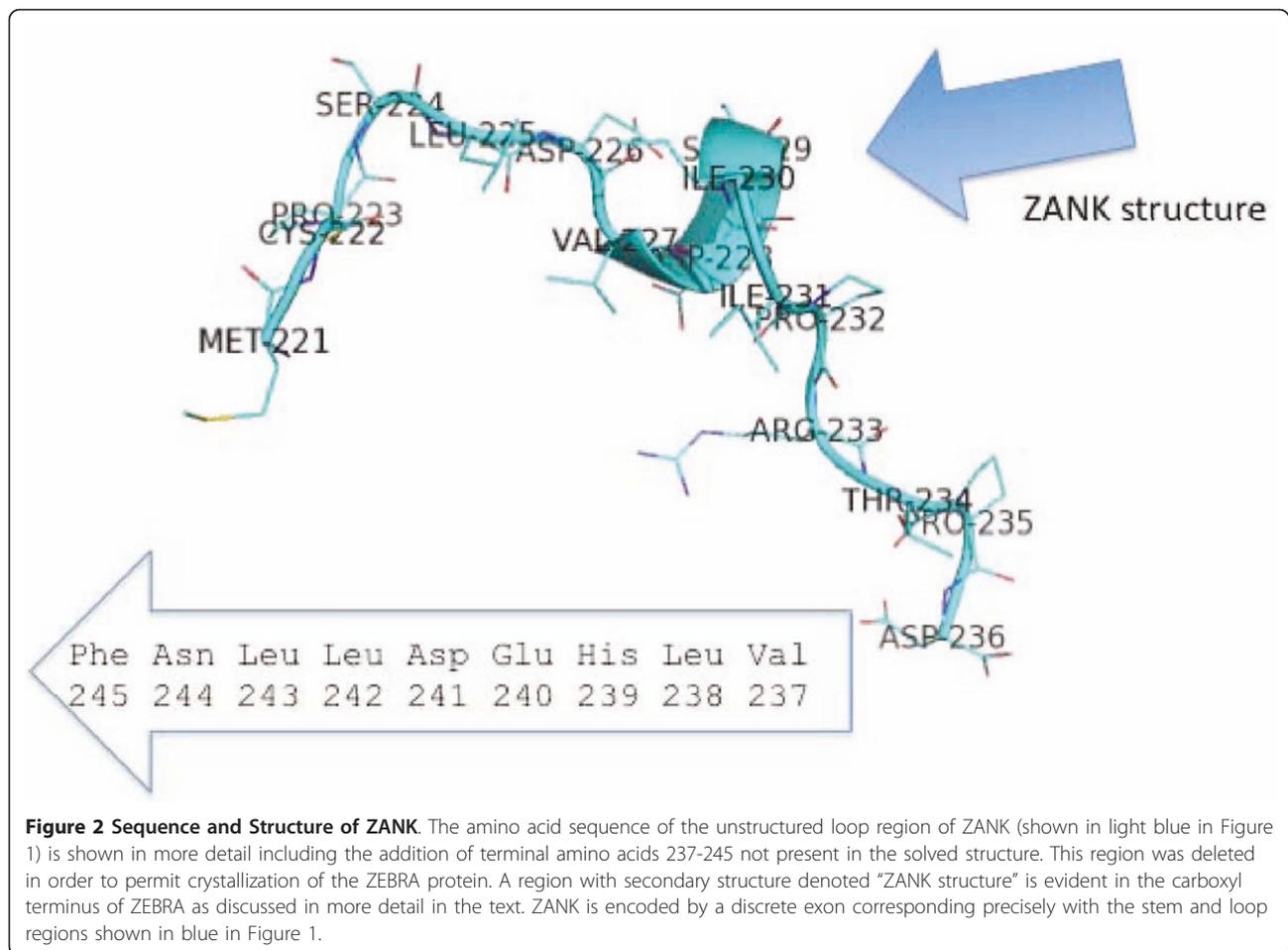
The ZANK region of ZEBRA was noted to be encoded by a single exon in all examined EBV strains coinciding exactly with the exon-intron junction of the ZEBRA protein third exon (Figure 2, 3). Thus, the p53 and NF- $\kappa B$  binding regions of ZEBRA, termed ZANK, were found to be separate both structurally and genetically from the DNA binding and trans-activation regions of ZEBRA (Figure 4).

### Alignment of ZANK with defined ankyrin proteins

ZEBRA protein and nucleotide sequences are as determined previously [26]. Individual ankyrin domains from published  $I\kappa B$  and related sequences obtained from genbank were identified using PyMOL and first aligned with ZANK and with each other using standard algorithms such as BLAST as illustrated in the text (Figure 5) Color scheme for Figure 5 can be found at [http://ekhidna.biocenter.helsinki.fi/pfam2/clustal\\_colours](http://ekhidna.biocenter.helsinki.fi/pfam2/clustal_colours)). Previously, some primary amino acid similarity between p53 binding ankyrins,  $I\kappa B$  ankyrins, invertebrate ankyrins and the ZEBRA carboxyl terminus have been illustrated and suggested to provide a basis for interactions between ZEBRA and ankyrin binding proteins [6,21].



**Figure 1 Structure of ZEBRA protein.** The carboxyl terminus of dimeric ZEBRA protein bound to a ZRE oligonucleotide (orange) is shown. This image was generated from previously determined ZEBRA molecular coordinates using PyMOL. The ZANK (ZEBRA ANKryn-like region) carboxyl terminus of ZEBRA is highlighted in blue. ZANK is composed of two regions, a highly structured alpha helix stem highlighted in darker blue (the ZANK stem region) and a partially solved unstructured region continuing to the end of the protein highlighted in light blue (the ZANK loop region).



Individual ankyrin domains in ASPP2 are denoted in this work based on previously reported optimized alignments between the p53BP2 fragment of ASPP2 and I $\kappa$ B. Thus, ank3 of p53BP2 corresponds to I $\kappa$ B ankyrin 3, p53BP2 ank4 corresponds to I $\kappa$ B ankyrin 4, and p53BP2 ank5 corresponds to I $\kappa$ B ankyrin 5 proceeding from the amino to carboxyl termini of both proteins respectively[6]. In a more recent naming convention, p53BP2 which was found to be a fragment of a larger protein known as ASPP2 (ankyrin repeat, SH3, and proline rich domain-containing protein number 2 also previously called apoptosis stimulating p53 binding protein 2) to indicate that it is a member of a multi-protein family. Using this nomenclature, p53BP2 ank3 corresponds to ASPP2 ankyrin 1, p53BP2 ank4 corresponds to ASPP2 ankyrin 2, and p53BP2 ank5 corresponds to ASPP2 ankyrin 3 [34].

#### Randomization analysis of ZANK and other defined ankyrin domains

A randomization shuffling strategy was developed to estimate the probability that observed amino acid similarities between ZANK, I $\kappa$ B and other related ankyrins

such as invertebrate cactus and unc-22 were not explained by the non-random amino acid composition of ankyrin proteins. The amino acid sequence of each ankyrin was randomized or shuffled and compared to every other possible sequence with the number of equal or better scores to non shuffled sequence shown. Repeated I $\kappa$ B and p53BP2 units of ankyrin are composed of a rigid alpha-helical stem and a less structured more variable loop region that may be under different evolutionary constraints. The borders of each ankyrin stem and loop were identified using PyMOL (Figure 5). Because of this bi-partite structure, analysis of the stem and loop regions was conducted both together and independently (Figure 6). Randomization of 10,000 of each stem and loop region was used in this analysis with correction for multiple comparisons. In this analysis, alignment of less than 500/10,000 randomized sequences is comparable to a P value of .05, shown in yellow. Alignment of less than 10/10,000 sequences comparable to a P value of .001 is shown in red.

Based on these studies the similarity between putative ZEBRA ankyrin stem and the cactus ankyrin stems 4 is

```

                203 204 205 206 207 208 209 210 211 212 213 214
TTTGTCCGCAG GTG GCT GCT GCC AAA TCA TCT GAA AAT GAC AGG CTG
      ^ Val Ala Ala Ala Lys Ser Ser Glu Asn Asp Arg Leu
                (Ser)
215 216 217 218 219 220 221 222 223 224 225 226 227 228 229
CGC CTC CTG TTG AAG CAG ATG TGC CCA AGC CTG GAT GTT GAC TCC
Arg Leu Leu Leu Lys Gln Met Cys Pro Ser Leu Asp Val Asp Ser

230 231 232 233 234 235 236 237 238 239 240 241 242 243 244
ATT ATC CCC CGG ACA CCA GAT GTT TTA CAC GAG GAT CTC TTA AAT
Ile Ile Pro Arg Thr Pro Asp Val Leu His Glu Asp Leu Leu Asn

245
TTC TCCCGTTATTGAAACCACGCCTGCTTCACG
Phe                                     ^

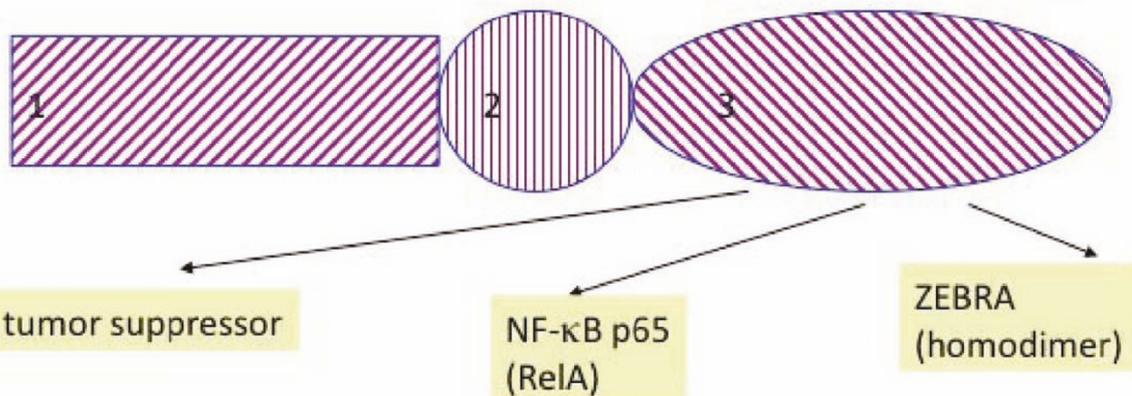
```

**Figure 3 ZANK region amino acids correspond precisely with exon 3 of BZLF-1 gene encoding ZEBRA.** DNA encoding the ZANK region and mRNA splice sites are shown. Using the same color scheme as in Figure 1, amino acids corresponding to the ZANK stem are shown in dark blue and those corresponding to the loop region are shown in light blue. DNA sequence is derived from the B-958 strain of EBV. A single amino acid polymorphism at aa 205 in this region has been identified resulting in a Ser substitution in certain strains of EBV such as Akata. Underlined sequences indicate regions of the putative ZANK loop present in the crystal structure. The splice points of ZEBRA exon 3 encoding ZANK is shown as “^” starting at the ZEBRA amino acid sequence “Val-Ala-Ala-Ala...” continuing to the end of the protein. The structured region of ZANK loop (Figure 1, 2) corresponds to ZANK sequences 227-230 Val-Asp-Ser-Ile. Residues His239 and Phe245 in the terminus of ZANK correspond to functionally conserved residues in the loops of other ankyrin proteins as discussed in more detail in the text.

**Domain 1:**  
Transcription  
activator

**Domain 2:**  
Site-specific  
DNA binding

**Domain 3:**  
Protein/Protein  
Interactions (ZANK)



**Figure 4 A schematic diagram of the ZEBRA protein illustrating the 3 functional domains of ZEBRA.** ZEBRA domains 1 and 2 contain the transcriptional activation and DNA binding regions of the protein, respectively, and are similar to Fos/Jun transcription factors. Domain 3, the carboxyl region of ZEBRA denoted ZANK in this work, is required for both dimerization of the ZEBRA protein as well as interactions with NF- $\kappa$ B and p53 transcription factors.



Anykrin and ZANK randomized and aligned  
 top: stems and loop, middle stem only, bottom loops only

Q/5	Zank	Cactus2	Cactus3	Cactus4	Cactus5	Ikbank2	Ikbank3	Ikbank4	Ikbank5	P53tp2ank3	P53tp2ank4	P53tp2ank5	Unc22
Zank	1.00E-04	7.48E-02	8.98E-01	1.16E-02	4.16E-02	7.89E-01	2.36E-02	6.63E-01	1.23E-01	6.02E-01	4.85E-01	4.16E-01	2.38E-01
Cactus2	1.00E-01	1.00E-04	7.50E-03	1.00E-04	2.60E-03	1.00E-04	7.10E-03	1.20E-03	1.00E-04	3.09E-01	1.00E-04	1.00E-04	1.77E-02
Cactus3	8.84E-01	8.60E-03	1.00E-04	2.03E-01	2.90E-01	3.80E-03	1.00E-04	7.09E-02	3.27E-02	5.71E-01	9.00E-04	2.74E-01	2.48E-01
Cactus4	8.40E-03	1.00E-04	1.95E-01	1.00E-04	2.40E-02	2.90E-03	4.38E-02	1.00E-04	1.00E-04	3.82E-01	1.00E-04	1.00E-04	2.00E-04
Cactus5	5.73E-02	1.10E-03	3.87E-01	2.54E-02	1.00E-04	1.80E-03	7.00E-04	3.83E-01	1.00E-04	5.86E-01	1.91E-02	3.50E-03	2.39E-01
Ikbank2	8.31E-01	1.00E-04	5.00E-03	4.01E-02	1.00E-03	1.00E-04	1.04E-02	2.90E-03	1.00E-04	4.72E-01	8.00E-04	3.00E-04	3.00E-04
Ikbank3	2.44E-02	6.20E-03	1.00E-04	4.01E-02	1.00E-03	7.50E-03	1.00E-04	8.43E-02	1.54E-01	5.49E-01	1.45E-01	4.64E-02	2.20E-01
Ikbank4	6.59E-01	4.00E-04	4.93E-02	1.00E-04	3.41E-01	1.80E-03	5.79E-02	1.00E-04	2.00E-04	2.54E-01	1.00E-04	7.80E-03	3.80E-03
Ikbank5	1.64E-01	2.00E-04	4.85E-02	1.00E-04	1.00E-04	1.00E-04	1.35E-01	3.01E-04	1.00E-04	2.77E-01	1.00E-04	1.00E-04	1.03E-02
P53tp2ank3	5.64E-01	3.01E-01	6.57E-01	2.82E-01	5.59E-01	4.44E-01	6.18E-01	1.48E-01	1.91E-01	1.00E-04	3.56E-01	9.00E-01	3.60E-01
P53tp2ank4	5.61E-01	1.00E-04	6.00E-04	1.00E-04	1.56E-02	9.00E-04	2.27E-01	1.00E-04	1.00E-04	4.41E-01	1.00E-04	2.00E-04	1.00E-04
P53tp2ank5	3.56E-01	1.00E-04	2.65E-01	1.00E-04	2.60E-03	3.00E-04	4.43E-02	2.06E-02	1.00E-04	9.17E-01	1.00E-04	1.00E-04	6.14E-02
Unc22	2.63E-01	6.50E-03	3.09E-01	2.00E-04	2.80E-01	3.00E-04	2.54E-01	1.50E-03	1.61E-02	4.00E-01	2.00E-04	8.72E-02	1.00E-04

Q/5	ZankA	Cactus2A	Cactus3A	Cactus4A	Cactus5A	Ikbank2A	Ikbank3A	Ikbank4A	Ikbank5A	P53tp2ank3A	P53tp2ank4A	P53tp2ank5A	Unc22A
ZankA	1.00E-04	2.80E-03	5.59E-01	1.40E-03	1.59E-01	5.55E-01	3.00E-01	6.26E-01	4.08E-02	2.86E-01	3.92E-01	2.37E-01	6.65E-02
Cactus2A	6.00E-03	1.00E-04	8.00E-04	1.00E-04	6.00E-04	1.00E-04	8.00E-04	9.00E-04	3.00E-04	4.01E-01	2.70E-03	2.00E-04	7.96E-02
Cactus3A	5.94E-01	2.00E-04	1.00E-04	8.65E-02	7.86E-02	6.00E-04	2.40E-03	3.35E-02	1.80E-03	3.10E-02	1.30E-03	2.28E-01	2.73E-01
Cactus4A	8.00E-03	1.00E-04	6.84E-02	1.00E-04	1.03E-02	8.08E-02	1.12E-02	1.00E-04	3.50E-03	1.34E-01	2.00E-04	1.90E-03	1.32E-02
Cactus5A	2.62E-01	1.00E-04	1.10E-01	3.36E-02	1.00E-04	4.00E-04	2.00E-04	1.46E-01	1.00E-04	5.63E-01	4.40E-03	5.00E-04	2.29E-02
Ikbank2A	8.00E-01	1.00E-04	8.00E-04	5.53E-02	2.00E-04	1.00E-04	1.00E-01	1.90E-03	2.00E-04	4.96E-01	1.20E-03	3.00E-04	4.00E-04
Ikbank3A	4.05E-01	2.00E-04	3.40E-03	1.81E-02	1.00E-04	8.00E-04	1.00E-04	1.43E-02	9.10E-03	2.83E-01	6.88E-02	9.30E-03	2.51E-01
Ikbank4A	7.32E-01	1.00E-04	1.45E-02	1.00E-04	1.12E-01	3.10E-03	9.00E-03	1.00E-04	1.10E-02	8.25E-02	2.00E-03	7.17E-02	3.62E-01
Ikbank5A	4.19E-02	6.00E-04	4.50E-03	2.70E-03	1.00E-04	1.00E-04	1.16E-02	1.11E-02	1.00E-04	6.11E-01	1.00E-04	4.50E-03	1.34E-01
P53tp2ank3A	3.95E-01	6.80E-01	7.16E-02	1.22E-01	5.24E-01	4.63E-01	3.12E-01	7.11E-02	6.49E-01	1.00E-04	1.46E-01	3.30E-01	5.10E-01
P53tp2ank4A	3.43E-01	9.00E-04	2.00E-03	2.00E-04	4.30E-03	1.10E-03	7.26E-02	3.10E-03	1.00E-04	1.55E-01	1.00E-04	2.00E-04	2.00E-02
P53tp2ank5A	2.57E-01	2.00E-04	2.61E-01	2.20E-03	1.00E-04	1.00E-04	7.00E-03	8.27E-02	1.20E-03	4.74E-01	1.00E-04	1.00E-04	4.40E-03
Unc22A	1.29E-01	4.00E-02	3.29E-01	9.40E-03	2.80E-02	1.00E-04	3.57E-01	3.36E-01	1.70E-01	6.28E-01	3.80E-02	7.50E-03	1.00E-04

Q/5	ZankB	Cactus2B	Cactus3B	Cactus4B	Cactus5B	Ikbank2B	Ikbank3B	Ikbank4B	Ikbank5B	P53tp2ank3B	P53tp2ank4B	P53tp2ank5B	Unc22B
ZankB	1.00E-04	6.16E-01	5.59E-01	6.88E-02	3.00E-04	6.50E-01	6.79E-02	1.91E-01	2.82E-01	3.43E-01	4.76E-01	7.20E-01	6.69E-01
Cactus2B	7.92E-01	1.00E-04	1.42E-01	2.78E-01	4.90E-01	1.00E-04	9.16E-01	4.86E-01	5.57E-01	2.17E-01	1.06E-01	2.70E-01	9.11E-02
Cactus3B	3.65E-01	1.44E-01	1.00E-04	1.83E-01	6.98E-02	1.63E-02	3.60E-03	6.31E-01	5.25E-01	3.28E-01	5.19E-01	6.27E-01	5.81E-01
Cactus4B	1.16E-01	3.08E-01	2.73E-01	1.00E-04	6.78E-01	3.54E-01	1.00E+00	1.00E-04	1.70E-03	6.55E-01	2.73E-02	1.72E-01	1.62E-02
Cactus5B	3.70E-03	4.78E-01	6.80E-02	6.09E-01	1.00E-04	4.12E-01	1.47E-01	1.00E+00	4.00E-04	8.47E-01	8.41E-02	6.14E-01	1.00E+00
Ikbank2B	8.56E-01	1.00E-04	1.92E-02	2.44E-01	5.02E-01	1.00E-04	4.50E-01	8.88E-01	4.48E-01	1.33E-01	8.76E-01	5.70E-01	2.27E-01
Ikbank3B	1.23E-01	9.33E-01	5.80E-03	1.00E+00	1.54E-01	3.44E-01	1.00E-04	4.39E-01	3.53E-01	4.59E-01	2.52E-01	4.74E-01	2.47E-01
Ikbank4B	2.54E-01	5.30E-01	5.64E-01	1.00E-04	1.00E+00	9.11E-01	5.41E-01	1.00E-04	1.35E-02	9.64E-01	9.40E-03	1.55E-01	2.45E-02
Ikbank5B	1.90E-01	5.28E-01	4.38E-01	1.90E-03	7.00E-04	4.24E-01	3.36E-01	1.10E-02	1.00E-04	8.37E-02	1.74E-02	3.30E-02	7.16E-02
P53tp2ank3B	4.19E-01	1.82E-01	2.56E-01	5.61E-01	8.14E-01	1.07E-01	5.47E-01	9.68E-01	6.62E-02	1.00E-04	9.14E-01	9.22E-01	1.99E-01
P53tp2ank4B	3.94E-01	3.35E-01	3.98E-01	2.52E-02	7.86E-02	8.59E-01	1.40E-01	8.90E-03	6.20E-03	9.28E-01	1.00E-04	8.54E-01	1.75E-01
P53tp2ank5B	7.94E-01	2.17E-01	1.40E-01	3.49E-01	3.87E-01	5.81E-01	7.41E-02	2.85E-02	9.62E-01	8.23E-01	1.00E-04	8.96E-01	8.96E-01
Unc22B	7.70E-01	2.36E-01	4.81E-01	3.50E-03	1.00E+00	2.16E-01	4.33E-01	5.70E-02	1.49E-01	1.48E-01	1.26E-01	7.52E-01	1.00E-04

**Figure 6 Statistical analysis of ZANK amino acid sequence similarity to other ankyrins.** Representative results of shuffling experiments for ZANK, vertebrate ankyrins, I $\kappa$ B $\alpha$ , p53BP2/ASPP2 ankyrins, and invertebrate ankyrins Cactus and Unc-22 are shown. Top: Randomizations of whole ankyrin repeats. Middle: ankyrin stem regions only. Bottom: ankyrin loop regions only. 10,000 randomly shuffled sequences of identical amino acid composition were generated in each cell. Shuffled sequences were aligned with indicated non-shuffled sequence and the number of scores greater than or equal to that of the non-shuffled input sequence is shown. Significant alignments ( $p < 0.05$ ) are shown highlighted in yellow while highly significant alignments ( $p < 0.001$ ) are shown highlighted in red.

from similarities between ZRE and AP-1 binding sites [8]. The structure of the carboxyl terminal region of the ZEBRA protein has not been determined in the published crystal structure as inclusion of this region prevented protein crystallization suggesting a disordered structure (Figure 1) [32]. This region of ZEBRA is also dispensable for DNA binding[11].

Efforts to discover a specific mechanism for protein-protein interactions of the carboxyl terminus region of ZEBRA based upon the partially solved structure have been unsuccessful, although some similarities are evident between this region and C/EBP- $\alpha$ , a CAAT-binding transcription factor whose dysregulation is implicated in acute myeloid leukemia [13]. Notably, the conserved carboxyl terminus of ZEBRA (Figure 2) diverges from DNA binding proteins of the BZIP family precisely at the intron border of exon 3 (Figure 3). Like the ZEBRA terminal exon, I $\kappa$ B binds to both p53 and NF- $\kappa$ B [6,21]. Exon 1 and 2 of ZEBRA thus encode a Fos/Jun homologue that activates DNA transcription to initiate the viral lytic program, while exon 3 (shared with trans-spliced

RAZ protein) provides both protein self-dimerization and binding to other master regulatory transcription factors such as p53 and NF- $\kappa$ B. These observations are analyzed in more detail in the remainder of this work.

**ZANK, the carboxyl terminal Zebra ANKYrin-like region**

Interactions between ZEBRA and NF- $\kappa$ B, p53 as well as other proteins can occur independently of ZEBRA DNA binding, since experiments *in vivo* have shown that in T-lymphocytes ZEBRA is not bound to DNA but localized to the cytoplasm where it blocks NF- $\kappa$ B translocation to the nucleus [6,21]. The full sequence of the carboxyl terminus exon (exon 3) of ZEBRA is shown in Figure 2, including both regions of the protein present in the crystal structure (blue highlights in Figure 1) and extreme carboxyl residues not present in the crystal structure [26]. We previously suggested that the carboxyl terminus of ZEBRA encodes an ankyrin-like structure responsible for the ability of ZEBRA to bind ankyrin binding proteins NF- $\kappa$ B and p53 [6,21]. For purposes of discussion, the carboxyl terminus of ZEBRA

encoded by exon 3 is termed ZANK (for ZEBRA ANKyrin-like region) in the remainder of this work. Remarkably, the crystal structure of ZEBRA protein (Figure 1) shows divergence from BZIP Fos/Jun DNA binding proteins precisely at the intron/exon borders of ZEBRA exon 3 (Figure 2, 3).

#### Comparison of ZANK with defined vertebrate and invertebrate ankyrin domains

ZANK spans part of the alpha helix of ZEBRA and unstructured regions at the carboxyl terminus of the protein (Figure 3, 4). The alpha-helical stem-like regions contribute most of the similarity between ZANK and individual ankyrins (Figure 5). This is evident in Figure 5 because most of the colored coded similar regions occur to the left of the figure in the stem regions (denoted s) rather than the loop regions (denoted l). This was confirmed by a randomization analysis (Figure 6). Randomizations demonstrate that the full ZANK sequence or loop region results in less significant alignments or no significant alignments for both ZANK and other ankyrins compared to randomization of the stem regions (Figure 6). The alpha helical regions of ankyrin binding proteins are highly structured to provide a rigid framework, while the loop regions are less structured allowing for binding to different substrates. Results are consistent with the hypothesis that ZANK encodes amino acids derived from a common ancestral ankyrin domain shared with I $\kappa$ B and p53 binding ankyrins and are most closely related to the invertebrate *Drosophila* ankyrin regulatory protein Cactus, rather than vertebrate ankyrins.

Statistical significance as highlighted in yellow in Figure 6 is defined as a given alignment found in less than 5% of shuffled sequences (less than 500 out of 10,000), corresponding to  $P < 0.05$  that the alignment would occur by chance. A more stringent definition of a non-random alignment (highlighted in red in Figure 6) is defined as a given alignment found in less than 0.1% of shuffled sequences (less than 10 out of 10,000). In all cases analyzed, including ZANK, the stem or alpha helix regions of ankyrins are more conserved across ankyrin protein families than the loop regions. For example, not all I $\kappa$ B $\alpha$  loops in host proteins are similar to each other, but all the I $\kappa$ B $\alpha$  stem regions are similar based upon 10,000 random shuffling of the respective sequences.

The diagonal of each randomization experiment shown is the identity of the sequence with itself shown in red, which as expected, occurs once per 10,000 shuffling experiments. Because each cell represents a characterization of an independent shuffling “experiment” the number of aligned sequences are slightly different even for identical comparisons as can be seen by comparing identical comparisons on either side of the diagonal

identity but because of the large number of shuffling per “experiment” results are usually similar for a given cell on either side of the diagonal.

#### ZANK is more similar to an invertebrate ankyrin domain from Cactus than vertebrate ankyrins

As shown in Figure 6, the ZANK stem was most similar to the ankyrin 4 stem region of invertebrate *Drosophila* Cactus regulatory protein at the highest level of significance (less than 10/10,000 of shuffled sequences). ZANK loop was most similar to invertebrate *Drosophila* Cactus regulatory protein ankyrin 5 loop region at the highest level of significance (less than 10/10,000 of shuffled sequences). Similarity was also evident between ZANK and the Cactus regulatory protein ankyrin 2 stem region and the I $\kappa$ B $\alpha$  ankyrin 3 and 5 stems as well as the whole protein. Similarity to both stem and loop regions of Cactus in addition to elements of I $\kappa$ B $\alpha$  is a surprising result since EBV is a vertebrate pathogen, while Cactus is an invertebrate host gene, suggesting that the divergence of ZANK from other ankyrins was ancient, possibly preceding divergence of p53 and NF- $\kappa$ B ankyrin binding proteins such as I $\kappa$ B $\alpha$ . Alternatively, a vertebrate virus encoding ZANK could have resulted from recombination with an invertebrate *Drosophila* virus since in some cases insect viruses can infect vertebrate cells. Existing phylogeny programs do not appear able to resolve these two possibilities as will be discussed in more detail elsewhere.

#### *In silico* Modeling of ZANK

In the remainder of this work we used PyMOL generated structures to model experimentally observed binding between ZEBRA protein and both NF- $\kappa$ B and p53 *in vitro* and *in vivo*. A limitation of this analysis is that the existing structure of ZANK is both partial and also solved in the absence of the partner ligands p53 or NF- $\kappa$ B. However, even with those limitations this analysis is sufficient to suggest a model of ZANK interactions with its partner ligands in which the ZANK loop adopts two different conformations. We have previously suggested that binding between ZEBRA carboxyl terminus and other proteins can be explained by the common descent of both NF- $\kappa$ B and p53 from an ancestral protein termed proto-p53/NF- $\kappa$ B. If this hypothesis is correct then it should be possible to identify a common structural signature or code between ZANK, NF- $\kappa$ B binding ankyrins and p53 binding ankyrins that ZANK has evolved to fit into. To use an analogy, if ankyrins resemble a key that can bind to their substrates resembling a lock, then the ZANK key should fit into both the NF- $\kappa$ B and p53 locks[6,21].

To explore the possibility of this “lock and key” mechanism, I $\kappa$ B $\alpha$  ankyrin structures were extracted and

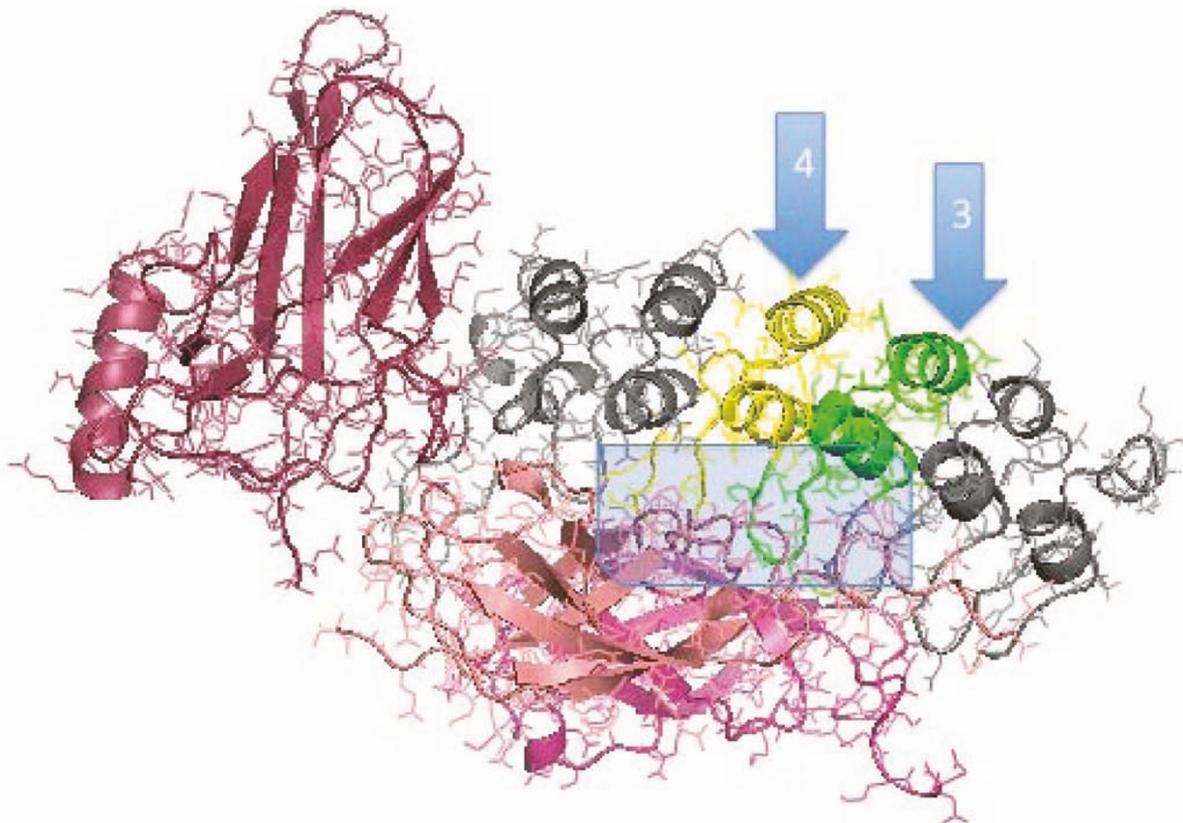
superimposed as shown, without any other alterations in the solved crystal structure coordinates. Since ZEBRA protein exists as a dimer of identical subunits (Figure 1), two adjacent ZANK loop domains in the dimerized ZEBRA protein could interact with adjacent I $\kappa$ B $\alpha$  ankyrin binding regions of NF- $\kappa$ B. As shown, Ank 3 and 4 of I $\kappa$ B $\alpha$  play a critical role in I $\kappa$ B $\alpha$  interactions with NF- $\kappa$ B (Figure 7). In support of this hypothesis, residues such as Histamine 239 in ZANK (Figure 3), not present in the crystal structure of ZEBRA due to instability in the absence of either co-ligand NF- $\kappa$ B and tumor suppressor p53, could align with a corresponding conserved stem histidine in I $\kappa$ B $\alpha$  Ank 3 and related ankyrin proteins (Figure 5). Binding to proteins such as NF- $\kappa$ B and p53 could stabilize ZANK and permit a crystal or structure to confirm this model.

**NF- $\kappa$ B binding ankyrins have an alternating key or code of short and long ankyrin loops suggesting a polymorphic ZANK loop domain**

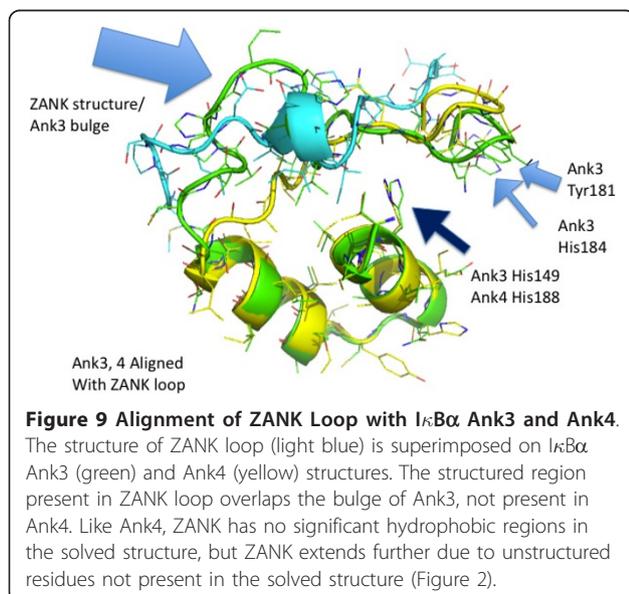
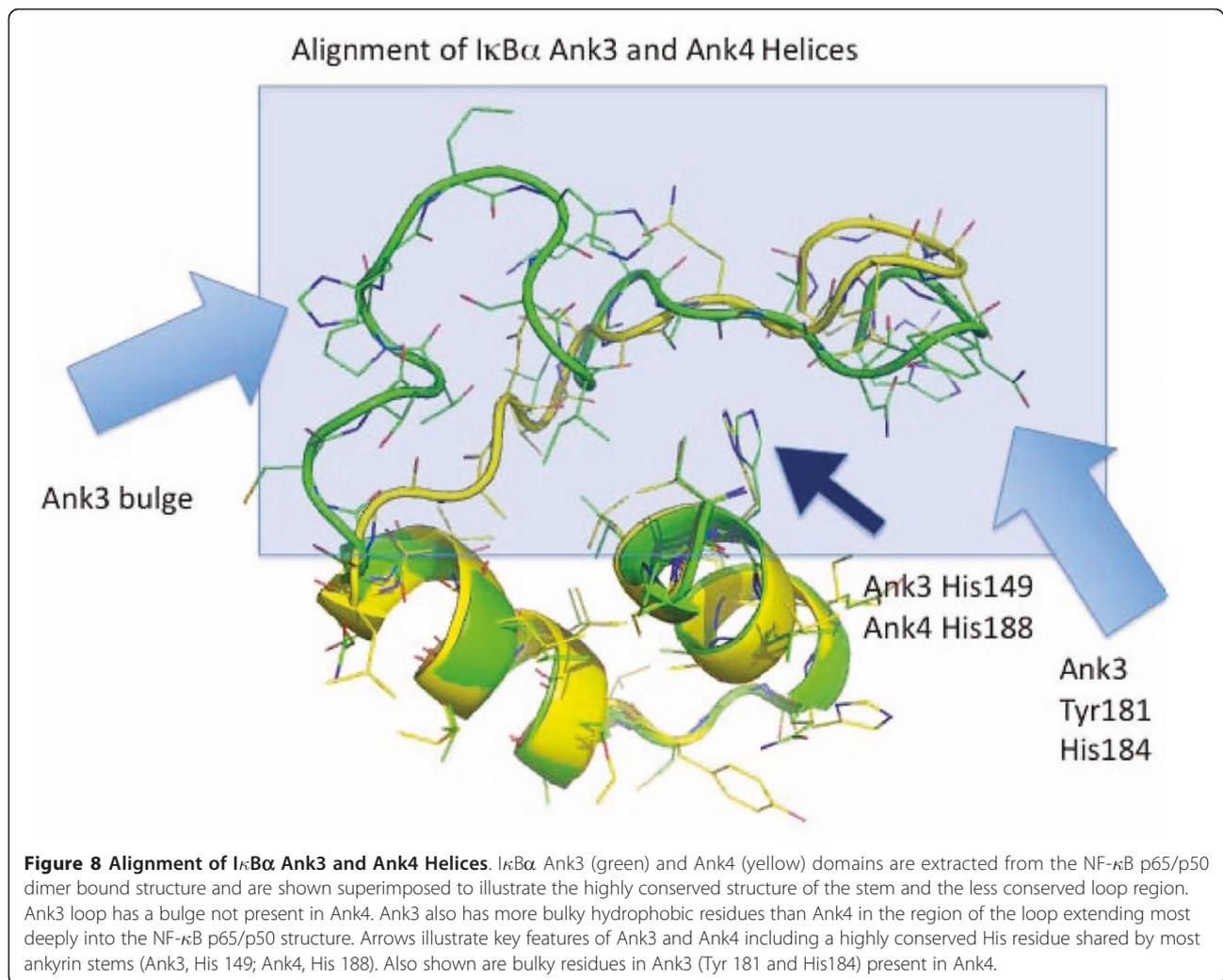
Ankyrin loop regions are variable in length while the stems are conserved in length (Figure 5). There is an

alternating pattern of short loops (approximately 15 amino acids between adjacent stems) and long loops (approximately 20 amino acids between adjacent stems) in all of the NF- $\kappa$ B binding ankyrins including vertebrate I $\kappa$ B $\alpha$  and invertebrate cactus. Thus both I $\kappa$ B $\alpha$  and cactus ankyrin loops 2, 4, and 6 are short and loops 1, 3, and 5 are long. To permit binding interactions with NF- $\kappa$ B, the ZANK loop region of ZANK would need to be dimorphic or sufficiently unstructured or flexible to conform to adjacent non-identical short and long ankyrin binding pockets in NF- $\kappa$ B, for example the regions of NF- $\kappa$ B binding to Ankyrins 3 and 4 of I $\kappa$ B $\alpha$  as shown (Figure 7, 8, 9). Because the ankyrin loops of I $\kappa$ B $\alpha$  ank5 are very similar in both primary and secondary structure to I $\kappa$ B $\alpha$  ank3, I $\kappa$ B $\alpha$  has an alternating pattern of Ank3 and Ank4-like subunits, and this alternating structure could permit multiple ZANK binding sites in a single NF- $\kappa$ B ligand (Figure 10).

Also resulting from this alternating short loop/long loop pattern, using PyMOL generated structures it is evident that I $\kappa$ B $\alpha$  ankyrin 3 has additional amino acids in its loop region forming a bulge not present in ankyrin



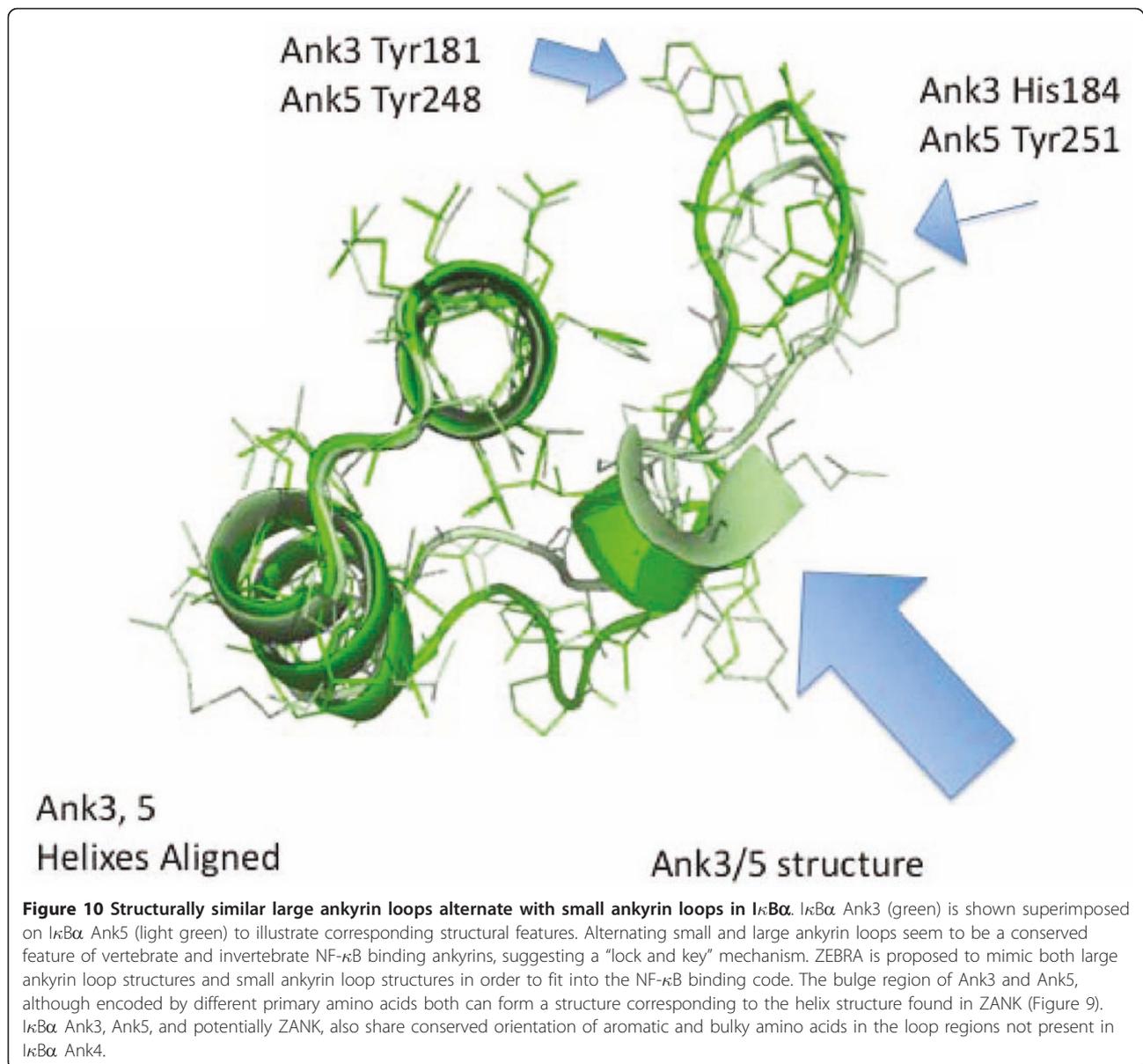
**Figure 7 Illustration of the Interaction of I $\kappa$ B $\alpha$  ankyrins with NF- $\kappa$ B.** NF- $\kappa$ B p65/p50 dimer subunits (red shadings) are shown bound to ankyrin repressor I $\kappa$ B $\alpha$ . I $\kappa$ B $\alpha$  ankyrin domains 3 (green) and 4 (yellow) are shown in color while other regions of the I $\kappa$ B $\alpha$  protein are shown in grey. Each ankyrin unit is composed of a structurally conserved alpha helix dimer backbone, and a less structurally conserved loop region inserting into the NF- $\kappa$ B p65/p50 groove.



4, and also has hydrophobic amino acids in its loop not present in ankyrin 4 (Figure 9). If ZEBRA were to bind to both  $\text{I}\kappa\text{B}\alpha$  ankyrin 3 and 4 binding regions in NF- $\kappa\text{B}$ , ZANK would need to adapt 2 different conformations to fit into the two different sites, for example by having free amino acids outside of the binding region at the Ankyrin 4 site, while utilizing additional amino acids in the extreme carboxyl region of the ZEBRA protein in the deeper Ankyrin 3 site. The bulge region present both ankyrin 3 and 5 loops is not present in ankyrin 4 loop (Figure 10), but corresponds to a structured region of the ZANK loop (Figure 9). In the  $\text{I}\kappa\text{B}\alpha$  structure, these structured regions of are proposed to form a  $\alpha$  helix like region stabilizing the  $\text{I}\kappa\text{B}\alpha$  loops.

#### Shared structural elements between p53 and NF- $\kappa\text{B}$ binding ankyrins identified as a potential binding mechanism of ZANK

We suggested previously that that the ability of ZEBRA to bind both NF- $\kappa\text{B}$  and p53 results from an

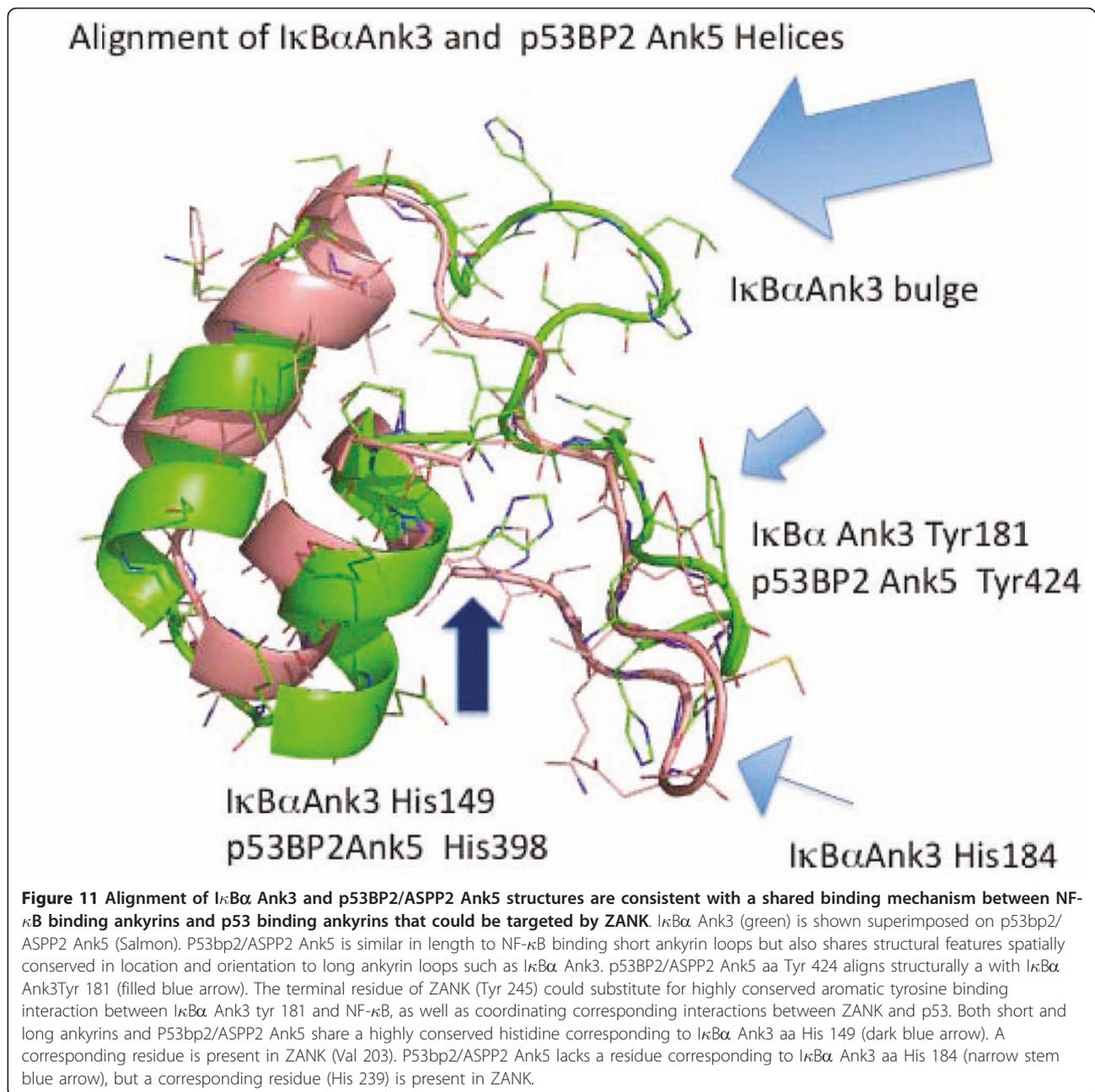


evolutionarily conserved binding pocket shared between the two proteins[6,21]. The solved structure of p53bp2 (more recently denoted ASPP2 (1YCS.pdb)) bound to p53 demonstrates that p53bp2/ASPP2 contacts p53 in only the extreme carboxyl terminus, ankyrins 4 and 5, of its ankyrin regions, although other binding modes could exist *in vivo*. Pymol generated alignment between I $\kappa$ B $\alpha$  ankyrin 3 and p53bp2/ASPP2 ankyrin 5 is shown (Figure 11). p53bp2/ASPP2 ankyrin 5 has a structure with features similar to I $\kappa$ B $\alpha$  ankyrin 3 – a highly conserved tyrosine is shared between ankyrin 3 and ASPP2 ankyrin 5, and this residue in I $\kappa$ B $\alpha$  is known to form a specific contact with a tyrosine in NF- $\kappa$ B p50/p65 dimer. Interestingly, the p53bp2/ASPP2 ankyrin loops

do not have a short long loop pattern (Figure 5) but instead have a length corresponding to the short I $\kappa$ B $\alpha$  ankyrin loops (approximately 15 aa), and a conserved tyrosine typical of the long loops. These observations are consistent with a shared binding signature between some NF- $\kappa$ B and p53 binding proteins that could be targeted by a polymorphic ZANK loop as a master key.

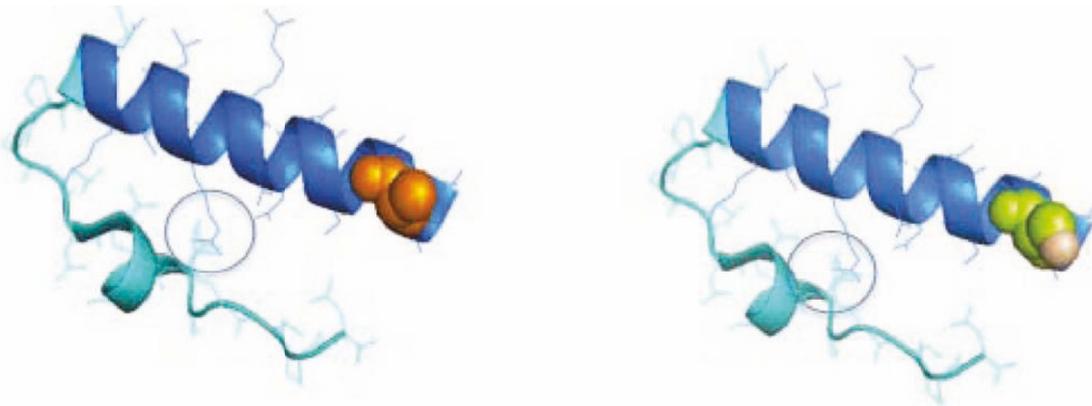
#### **Identification of a cancer associated sequence variant in ZANK potentially altering p53binding interactions**

An advantage of PyMOL generated *in silico* structural analysis of ankyrin domains in NF- $\kappa$ B proteins and p53 binding proteins is that it may now be possible to predict the behavior of polymorphisms in ZEBRA that increase



or decrease the relative strength of binding to NF- $\kappa B$  versus p53 proteins (Figure 12). Alterations in ZANK amino sequences could alter the ability of ZEBRA to transiently block p53 function during viral replication and thus alter viral growth properties, accounting in part for the recently described association of ZEBRA with increased mutagenic and carcinogenic properties in a murine model[14]. For example, the limited sequence polymorphism data available for the ZANK region of ZEBRA currently suggests that a mutation in tumor associated ZEBRA from Akata cells is different from canonical B-958 ZEBRA at a single amino acid in the

ZANK region (ZEBRA exon 3). Remarkably, primary sequence comparison between ZANK,  $I\kappa B$  and related proteins and ASPP2 ankyrins reveals that the substitution of a serine for an alanine at amino acid 205 (Figure 3) is not random but rather converts ZANK from an  $I\kappa B$ -like AAA motif to a p53BP2/ASPP2-like ASA motif, also shared with a Cactus ankyrin (Figure 5). Unlike alanine, a serine residue may be covalently modified *in vivo* by phosphorylation, further extending the possible functional consequences of this polymorphism[35]. These observations suggest further study of ZANK polymorphisms *in vivo* and their effects on NF- $\kappa B$  and p53 binding



Left: ZANK with B-958 residue alanine 205 (orange) indicated  
Right: ZANK with Akata residue serine 205 (serine green with pink hydroxyl).  
Position of contact between ZANK stem and loop residues indicated (circle)

**Figure 12 ZANK Structure with S205A polymorphism illustrated.** A functionally conserved polymorphism in ZANK, A205 is altered to Ser is present in some EBV positive lymphoma cell lines such as Akata. A similar serine rather than alanine residue is present in some p53 binding protein ankyrin domains but not in NF- $\kappa$ B binding ankyrin domains. Changes in the stem or loop regions of ZANK could alter ZANK loop interactions with p53 and NF- $\kappa$ B particularly if covalently modified by serine phosphorylation. The circled region indicates a potential interaction between ZANK stem and loop amino acids.

*in vitro* could identify associations between certain EBV strain polymorphisms and cancer phenotypes.

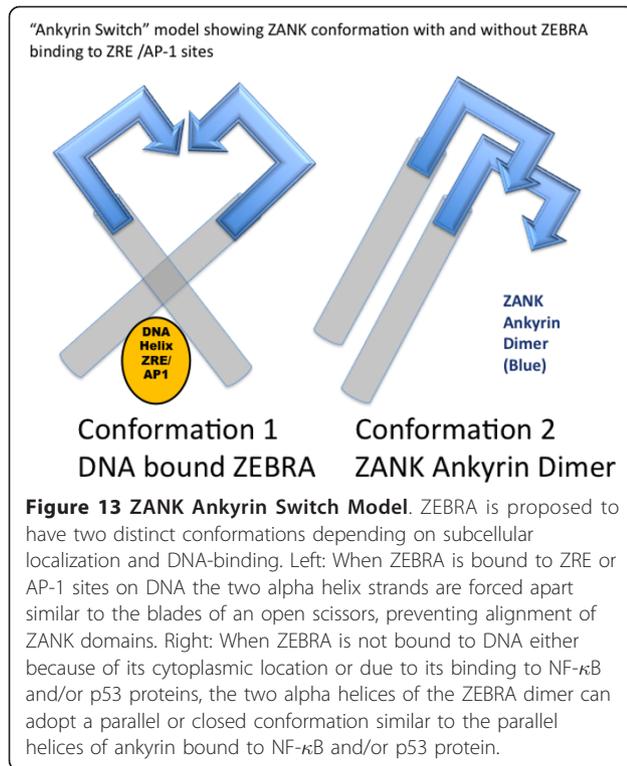
#### **A model of antagonistic interactions between ZANK NF- $\kappa$ B and p53 protein binding and ZEBRA fos/jun DNA binding**

As outlined in this manuscript, the extreme carboxyl terminal exon of ZEBRA encodes an amino acid sequence termed ZANK that is functionally and evolutionarily distinct from DNA binding regions of ZEBRA protein. The distinct origin and structure of ZANK suggest that ZANK domain could have a regulatory role on the DNA binding regions of the ZEBRA protein in addition to cytoplasmic effects on NF- $\kappa$ B and p53 proteins (Figure 1). In particular, ZANK could modulate ZEBRA DNA binding through conformational changes in the Fos/Jun like region of ZEBRA (Figure 4) triggered by binding between ZANK and proteins of the NF- $\kappa$ B and p53 family and *vice versa*, modulating the transcriptional activity of the cellular transcription machinery (Figure 13).

Binding of ZANK to proteins of the NF- $\kappa$ B and p53 families could stimulate, inhibit or have no effect of ZEBRA binding to DNA. As shown in Figure 13, structural considerations based on the partial structure of ZEBRA and ZANK support an antagonistic role between

ZEBRA DNA binding and ZANK interactions with NF- $\kappa$ B and p53 proteins. Based upon the crystal structure of ZEBRA bound to a ZRE/AP1 DNA binding oligonucleotide (Figure 1), it is apparent that DNA bound to ZEBRA causes the two DNA binding alpha helices to move away from each other. The movement of the ZANK regions at the termini of the helices in a region not bound to DNA is analogous to the blades of a scissor in the open position. A parallel orientation of the ZANK stems, similar to the closed blades of a scissor is required to form an ankyrin-like dimer capable of binding adjacent regions of NF- $\kappa$ B and p53 proteins (Figure 7).

Thus, ZANK could function as a NF- $\kappa$ B and p53 protein dependent switch in the nucleus of EBV infected B-lymphocytes, turning ZEBRA DNA binding off in the presence of high levels of nuclear NF- $\kappa$ B and p53, and conversely activating ZEBRA DNA binding in the absence of nuclear NF- $\kappa$ B and p53 protein (Figure 13). In physiologic terms, these observations are consistent with high levels of active NF- $\kappa$ B and p53 proteins present in the nucleus of latent B-lymphocytes during viral latency, which during viral latency would tend to block ZEBRA DNA binding and prevent activation of viral lytic growth. Conversely, transiently decreased levels of nuclear NF- $\kappa$ B proteins in latently infected B-



lymphocytes, due for example to cell stress or starvation could permit increased ZEBRA DNA binding and activation of viral lytic growth, coupling levels of cytoplasmic growth signals mediated by NF- $\kappa$ B and p53 proteins to the viral life cycle.

Later in the viral lytic replication cycle, the NF- $\kappa$ B and p53 protein binding function of ZANK could further trigger cellular apoptosis and facilitate infectious viral release directly through cytoplasmic effects on NF- $\kappa$ B and p53 protein. In T lymphocytes and possibly other non B-lymphocyte cells infected by EBV but incapable of supporting stable viral latency, ZEBRA is almost entirely present in the cell cytoplasm. In these cells, ZEBRA could initially inhibit NF- $\kappa$ B and cause activation of p53 protein and cellular apoptosis as a mechanism of viral immunosuppression.

Transplanted RAZ, containing ZANK but not capable of ZEBRA DNA binding would then serve as an amplifier of the NF- $\kappa$ B and p53 protein functions of ZEBRA, and a corresponding RAZ-like function might explain the non-DNA binding functions of a ZEBRA homologue in HSV-8, the etiologic agent of Kaposi's sarcoma and other human cancers. These possibilities could in principle be tested *in vitro* and *in vivo* through ZEBRA with ZANK domains engineered for either increased or reduced effects on NF- $\kappa$ B and/or p53 protein through mutation of conserved ZANK amino acids interacting

with NF- $\kappa$ B and p53 protein identified with the structural models and analysis presented in this work.

## Conclusions

Characterization of B cell lympho-proliferative disorders has demonstrated that EBV viral latency proteins can transform and immortalize B-lymphocytes *in vitro* and *in vivo*. A causal role of the virus in carcinogenesis is suggested in some tumors because the viral genome sometimes remains in the malignant cell. However, the mechanism of viral carcinogenesis cannot simply be a function of cellular immortalization by the virus since although the vast majority of humans harbor EBV immortalized B lymphocytes, EBV-genome associated cancers are rare [36]. EBV infection may thus also be associated with carcinogenesis not only through cellular immortalization but also through transient mechanisms including destabilization of host genomes [37-40].

A novel role of ZEBRA and ZANK in viral carcinogenesis is suggested in this work. ZEBRA plays a critical role in the viral latent to lytic switch, possibly in part through response to NF- $\kappa$ B and p53 activities by ZANK (Figure 13). Alternating cycles of partial lytic growth triggered by ZEBRA activation and ZEBRA DNA binding followed by ZANK interactions with NF- $\kappa$ B and p53 could synergize with genomic instability contributed by cellular or viral recombinase activation in EBV infected cells. Partial cycles of viral activation and inactivation without completion of the viral life cycle could be particularly relevant in the "hit and run" scenario, proposed for viral carcinogenesis and immune dysregulation as proposed in epithelial breast cancers [41-47].

Notably, it is possible that viral inactivation of p53 originates indirectly or as an evolutionary "spandrel." A spandrel refers to a new phenotype generated by alteration of a related developmental pathway. Because of the evolutionary relationship between NF- $\kappa$ B and p53 through descent from a common ancestral transcription factor, an unavoidable consequence of the inactivation of NF- $\kappa$ B by an ankyrin-like region of ZEBRA is a spandrel-like transient inactivation of p53. Restated, the common origins of NF- $\kappa$ B and p53 conserve a similar ankyrin binding region in both proteins, so any viral protein that targets this site on NF- $\kappa$ B proteins could, in parallel, affect p53 generating new p53 dependent phenotypes. This may in part contribute to the selective loss of the I $\kappa$ B $\alpha$  locus in tumors, a conserved feature of several EBV associated tumors [48,49]. The ZANK spandrel effects on p53 are maintained because of a selective benefit to the virus from related transient inactivation of the NF- $\kappa$ B transcription family and the innate immune system, as well as effects of ZANK on regulation of ZEBRA DNA binding.

Mutant p53 proteins were first characterized as oncogenes because their presence was associated with tumor phenotypes including tumor growth and resistance to apoptosis[50-52]. Surprisingly, wild type p53 was later shown to be a tumor suppressor through the effects of the protein as a DNA-binding transcription factor, rather than a tumor promoter. These contradictory effects could result from opposing tumor promoting spandrel effects between numerous viral proteins such as ZEBRA and mutant p53 proteins, versus tumor suppressing DNA binding gene activation by wild type p53 protein. Thus ZANK might augment the effects of mutant p53, not directly related to p53 DNA binding, but rather due to effects on structurally similar regions of NF- $\kappa$ B proteins required for cellular proliferation.

From an evolutionary perspective, it is likely that EBV encoded ZEBRA protein and a related protein in Human Herpes Virus 8 diverged from other BZIP DNA binding proteins such as Fos/Jun through viral capture of a terminal exon encoding a single ankyrin-like stem and loop. This domain was capable of dimerization into a structure resembling adjacent I $\kappa$ B ankyrins, and was subject to positive selection through interactions with NF- $\kappa$ B proteins. However, because of the shared ankyrin binding domain between NF- $\kappa$ B and p53 proteins an unfortunate consequence of NF- $\kappa$ B inhibition by ZANK would also unavoidably be inactivation of p53. This could be particularly important as a pathogenic mechanism of EBV-related genomic instability in epithelial cells and T lymphocytes that do not support the establishment of viral latency or replication but can sustain a limited form a viral lytic growth that includes high levels of expression of a trans-spliced protein "RAZ" containing ZANK, but without the DNA-binding transcriptional activity on ZRE. In this context, these cells are "dead end hosts" in their inability to support viral replication but can nevertheless suffer the pathophysiological and molecular consequences of viral infection.

Viral pre-conditioning by common viral pathogens has been indirectly implicated in the clustering of human cancers both geographically and temporally based upon indirect evidence in human cancers that defective or abortive replication of common viral pathogens can account for clustering of cancers[53]. In the "viral conditioning" scenario, abortive cycles of viral replication can promote genomic instability. A model is presented here for the specific case of EBV. In the case of EBV and, potentially, other common viruses that interact with both NF- $\kappa$ B and p53, transient lytic gene expression in both B lymphocytes and non-B cells could provide a mechanism for this viral pre-conditioning where defective or aborted viral replication contributes to malignancy as supported by indirect epidemiologic evidence [54-56]. These effects of ZANK could be of particular

importance not only in B lymphocytes but in epithelial and T cell infection in which the virus expresses lytic gene products but in most cases does not establish latency[57]. In these cell types transient inactivation of the tumor suppressor p53 would occur in the context of activation of endogenous somatic recombination pathways, potentially resulting in re-arrangement, instability and mutation of endogenous oncogenes.

In summary, we have demonstrated that ZEBRA is a hybrid protein resulting from viral capture of both a Fos/Jun like transcription factor and a polymorphic I $\kappa$ B-like ankyrin carboxyl terminus. Primary sequence and structural similarities are evident between the carboxyl terminus of ZEBRA protein and the conserved stem regions of I $\kappa$ B $\alpha$  and p53BP2/ASPP2 ankyrins. It is suggested that a shared binding pocket in NF- $\kappa$ B is in turn bound by all of these ankyrin proteins and that ZANK is analogous to a "master key" capable of fitting into both NF- $\kappa$ B and p53 protein "locks". In the case of ZEBRA, protein binding to NF- $\kappa$ B provides a selective advantage to the virus through inactivation of the immune response against the virus during viral lytic replication, simultaneously and unavoidably altering the activation of p53.

#### Acknowledgements

LYG would like to thank Dr. Leslie Krueger for helpful discussions regarding EBV "hit-and-run" hypothesis and Dr. David H. Wagner for conversations and ideas regarding NF-kappa B and other topics including autoimmunity. Dr. Dreyfus (DHD) has been supported in the past by NIH training grants T32-GM07288 and T32-AI07365. DHD acknowledges helpful discussions with Dr. Hans Helmut Niller, University of Regensburg, Germany, Jill Countryman, Yale School of Medicine, New Haven CT, and Dr. Cedric Feschote, University of Texas, Austin as well as members of the Weizmann Institute in Israel (in alphabetical order) Drs. Boris Brumshtein, Moshe Oren, Zippi Shakked and Joel Sussman. DHD and LYG are co-founders of a Keren Pharmaceutical Inc., San Mateo, CA. DHD is Assistant Clinical Professor, Department of Pediatrics, Yale School of Medicine, New Haven, CT and a member of the Center for Allergy, Asthma, and Clinical Immunology (CAAI), Waterbury, CT, and JC and YL are associated with Yale University, however none of the observations, proposals and/or conclusions of this work are related to the operations of Keren, Yale, or CAAI.

#### Author details

<sup>1</sup>Department of Pediatrics, Yale SOM, 488 Norton Parkway, New Haven CT 06511, USA. <sup>2</sup>Department of Statistics, Yale University, P.O. Box 208290, New Haven CT 06520-8290, USA. <sup>3</sup>Keren Pharmaceutical, 63 Bovet Rd #143, San Mateo, CA 94402-3104, USA.

#### Authors' contributions

DHD designed and executed experiments confirming interactions between ZEBRA, p53, and NF- $\kappa$ B, identified similarities between ZEBRA and ankyrin amino acids, conceived and executed Pymol models and other artwork, drafted manuscript, and paid costs of manuscript processing. YL, JTC conceived and executed statistical analysis and alignments of ZANK and related ankyrins, and edited the manuscript. LYG provided historical background and terminology of previous research in viral carcinogenesis, p53, and NF- $\kappa$ B, and ankyrin proteins, designed and executed experiments confirming interactions between ZEBRA, p53, and NF- $\kappa$ B, identified similarities between ZEBRA and ankyrin amino acids, and edited the manuscript and artwork. All authors read and approved the final manuscript.

### Competing interests

The authors declare that they have no competing interests.

Received: 3 March 2011 Accepted: 5 September 2011

Published: 5 September 2011

### References

1. Niller HH, Salamon D, Ilg K, Koroknai A, Banati F, Schwarzmann F, Wolf H, Minarovits J: **EBV-associated neoplasms: alternative pathogenetic pathways.** *Med Hypotheses* 2004, **62**:387-391.
2. Niller HH, Salamon D, Rahmann S, Ilg K, Koroknai A, Banati F, Schwarzmann F, Wolf H, Minarovits J: **A 30 kb region of the Epstein-Barr virus genome is colinear with the rearranged human immunoglobulin gene loci: implications for a "ping-pong evolution" model for persisting viruses and their hosts. A review.** *Acta Microbiol Immunol Hung* 2004, **51**:469-484.
3. Dreyfus DH, Kelleher CA, Jones JF, Gelfand EW: **Epstein-Barr virus infection of T cells: implications for altered T-lymphocyte activation, repertoire development and autoimmunity.** *Immunol Rev* 1996, **152**:89-110.
4. Dreyfus DH: **Immune system: success owed to a virus?** *Science* 2009, **325**:392-393.
5. Dreyfus DH: **Paleo-immunology: evidence consistent with insertion of a primordial herpes virus-like element in the origins of acquired immunity.** *PLoS One* 2009, **4**:e5778.
6. Dreyfus DH, Nagasawa M, Gelfand EW, Ghoda LY: **Modulation of p53 activity by I $\kappa$ B $\alpha$ : evidence suggesting a common phylogeny between NF- $\kappa$ B and p53 transcription factors.** *BMC Immunol* 2005, **6**:12.
7. Miller G, El-Guindy A, Countryman J, Ye J, Gradoville L: **Lytic cycle switches of oncogenic human gammaherpesviruses.** *Adv Cancer Res* 2007, **97**:81-109.
8. Petosa C, Morand P, Baudin F, Moulin M, Artero JB, Muller CW: **Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein.** *Mol Cell* 2006, **21**:565-572.
9. Gutsch DE, Holley-Guthrie EA, Zhang Q, Stein B, Blanan MA, Baldwin AS, Kenney SC: **The bZIP transactivator of Epstein-Barr virus, BZLF1, functionally and physically interacts with the p65 subunit of NF- $\kappa$ B.** *Mol Cell Biol* 1994, **14**:1939-1948.
10. Morrison TE, Kenney SC: **BZLF1, an Epstein-Barr virus immediate-early protein, induces p65 nuclear translocation while inhibiting p65 transcriptional function.** *Virology* 2004, **328**:219-232.
11. Schelcher C, Al Mehairi S, Verrall E, Hope Q, Flower K, Bromley B, Woolfson DN, West MJ, Sinclair AJ: **Atypical bZIP domain of viral transcription factor contributes to stability of dimer formation and transcriptional function.** *J Virol* 2007, **81**:7149-7155.
12. Schelcher C, Valencia S, Delecluse HJ, Hicks M, Sinclair AJ: **Mutation of a single amino acid residue in the basic region of the Epstein-Barr virus (EBV) lytic cycle switch protein Zta (BZLF1) prevents reactivation of EBV from latency.** *J Virol* 2005, **79**:13822-13828.
13. McDonald CM, Petosa C, Farrell PJ: **Interaction of Epstein-Barr virus BZLF1 C-terminal tail structure and core zipper is required for DNA replication but not for promoter transactivation.** *J Virol* 2009, **83**:3397-3401.
14. Ma SD, Hegde S, Young KH, Sullivan R, Rajesh D, Zhou Y, Jankowska-Gan E, Burlingham WJ, Sun X, Gullely ML, et al: **A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas.** *J Virol* 2011, **85**:165-177.
15. Morrison TE, Mauser A, Wong A, Ting JP, Kenney SC: **Inhibition of IFN- $\gamma$  signaling by an Epstein-Barr virus immediate-early protein.** *Immunity* 2001, **15**:787-799.
16. Morrison TE, Mauser A, Klingelhut A, Kenney SC: **Epstein-Barr virus immediate-early protein BZLF1 inhibits tumor necrosis factor alpha-induced signaling and apoptosis by downregulating tumor necrosis factor receptor 1.** *J Virol* 2004, **78**:544-549.
17. Huxford T, Huang DB, Malek S, Ghosh G: **The crystal structure of the I $\kappa$ B $\alpha$ /NF- $\kappa$ B complex reveals mechanisms of NF- $\kappa$ B inactivation.** *Cell* 1998, **95**:759-770.
18. Jacobs MD, Harrison SC: **Structure of an I $\kappa$ B $\alpha$ /NF- $\kappa$ B complex.** *Cell* 1998, **95**:749-758.
19. Kenney JL, Guinness ME, Curiel T, Lacy J: **Antisense to the Epstein-Barr virus (EBV)-encoded latent membrane protein 1 (LMP-1) suppresses LMP-1 and bcl-2 expression and promotes apoptosis in EBV-immortalized B cells.** *Blood* 1998, **92**:1721-1727.
20. Feng WH, Westphal E, Mauser A, Raab-Traub N, Gullely ML, Busson P, Kenney SC: **Use of adenovirus vectors expressing Epstein-Barr virus (EBV) immediate-early protein BZLF1 or BRLF1 to treat EBV-positive tumors.** *J Virol* 2002, **76**:10951-10959.
21. Dreyfus DH, Nagasawa M, Pratt JC, Kelleher CA, Gelfand EW: **Inactivation of NF- $\kappa$ B by EBV BZLF1-encoded ZEBRA protein in human T cells.** *J Immunol* 1999, **163**:6261-6268.
22. Kelleher CA, Paterson RK, Dreyfus DH, Streib JE, Xu JW, Takase K, Jones JF, Gelfand EW: **Epstein-Barr virus replicative gene transcription during de novo infection of human thymocytes: simultaneous early expression of BZLF1 and its repressor RAZ.** *Virology* 1995, **208**:685-695.
23. Sinclair AJ: **bZIP proteins of human gammaherpesviruses.** *J Gen Virol* 2003, **84**:1941-1949.
24. Sinclair AJ: **Unexpected structure of Epstein-Barr virus lytic cycle activator Zta.** *Trends Microbiol* 2006, **14**:289-291.
25. Al Mehairi S, Cerasoli E, Sinclair AJ: **Investigation of the multimerization region of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) protein K-bZIP: the proposed leucine zipper region encodes a multimerization domain with an unusual structure.** *J Virol* 2005, **79**:7905-7910.
26. Packham G, Brimmell M, Cook D, Sinclair AJ, Farrell PJ: **Strain variation in Epstein-Barr virus immediate early genes.** *Virology* 1993, **192**:541-550.
27. Furnari FB, Zacny V, Quinlivan EB, Kenney S, Pagano JS: **RAZ, an Epstein-Barr virus transdominant repressor that modulates the viral reactivation mechanism.** *J Virol* 1994, **68**:1827-1836.
28. Zhang Q, Gutsch D, Kenney S: **Functional and physical interaction between p53 and BZLF1: implications for Epstein-Barr virus latency.** *Mol Cell Biol* 1994, **14**:1929-1938.
29. Mauser A, Saito S, Appella E, Anderson CW, Seaman WT, Kenney S: **The Epstein-Barr virus immediate-early protein BZLF1 regulates p53 function through multiple mechanisms.** *J Virol* 2002, **76**:12503-12512.
30. Dreyfus DH, Nagasawa M, Kelleher CA, Gelfand EW: **Stable expression of Epstein-Barr virus BZLF1-encoded ZEBRA protein activates p53-dependent transcription in human Jurkat T-lymphoblastoid cells.** *Blood* 2000, **96**:625-634.
31. Bailey SG, Verrall E, Schelcher C, Rhie A, Doherty AJ, Sinclair AJ: **Functional interaction between Epstein-Barr virus replication protein Zta and host DNA damage response protein 53BP1.** *J Virol* 2009, **83**:11116-11122.
32. Hicks MR, Balesaria S, Medina-Palazon C, Pandya MJ, Woolfson DN, Sinclair AJ: **Biophysical analysis of natural variants of the multimerization region of Epstein-Barr virus lytic-switch protein BZLF1.** *J Virol* 2001, **75**:5381-5384.
33. Benyamini H, Leonov H, Rotem S, Katz C, Arkin IT, Friedler A: **A model for the interaction between NF- $\kappa$ B and ASP2 suggests an I $\kappa$ B-like binding mechanism.** *Proteins* 2009, **77**:602-611.
34. Ahn J, Byeon IJ, Byeon CH, Gronenborn AM: **Insight into the structural basis of pro- and antiapoptotic p53 modulation by ASPP proteins.** *J Biol Chem* 2009, **284**:13812-13822.
35. El-Guindy AS, Paek SY, Countryman J, Miller G: **Identification of constitutive phosphorylation sites on the Epstein-Barr virus ZEBRA protein.** *J Biol Chem* 2006, **281**:3085-3095.
36. Thorley-Lawson DA: **EBV the prototypical human tumor virus—just how bad is it?** *J Allergy Clin Immunol* 2005, **116**:251-261, quiz 262.
37. Niller HH, Salamon D, Uhlig J, Ranf S, Granz M, Schwarzmann F, Wolf H, Minarovits J: **Nucleoprotein structure of immediate-early promoters Zp and Rp and of oriLyt of latent Epstein-Barr virus genomes.** *J Virol* 2002, **76**:4113-4118.
38. Niller HH, Wolf H, Minarovits J: **Epigenetic dysregulation of the host cell genome in Epstein-Barr virus-associated neoplasia.** *Semin Cancer Biol* 2009, **19**:158-164.
39. Takacs M, Banati F, Koroknai A, Segesdi J, Salamon D, Wolf H, Niller HH, Minarovits J: **Epigenetic regulation of latent Epstein-Barr virus promoters.** *Biochim Biophys Acta* 2010, **1799**:228-235.
40. Li MJ, Maizels N: **Activation and targeting of immunoglobulin switch recombination by activities induced by EBV infection.** *J Immunol* 1999, **163**:6659-6664.
41. Perkins RS, Sahn K, Marando C, Dickson-Witmer D, Pahnke GR, Mitchell M, Petrelli NJ, Berkowitz IM, Soteropoulos P, Aris VM, et al: **Analysis of Epstein-**

- Barr virus reservoirs in paired blood and breast cancer primary biopsy specimens by real time PCR. *Breast Cancer Res* 2006, **8**:R70.
42. Yasui Y, Potter JD, Stanford JL, Rossing MA, Winget MD, Bronner M, Daling J: **Breast cancer risk and "delayed" primary Epstein-Barr virus infection.** *Cancer Epidemiol Biomarkers Prev* 2001, **10**:9-16.
  43. Arbach H, Viglasky V, Lefeu F, Guinebretiere JM, Ramirez V, Bride N, Boualaga N, Bauchet T, Peyrat JP, Mathieu MC, *et al*: **Epstein-Barr virus (EBV) genome and expression in breast cancer tissue: effect of EBV infection of breast cancer cells on resistance to paclitaxel (Taxol).** *J Virol* 2006, **80**:845-853.
  44. Huang J, Chen H, Hutt-Fletcher L, Ambinder RF, Hayward SD: **Lytic viral replication as a contributor to the detection of Epstein-Barr virus in breast cancer.** *J Virol* 2003, **77**:13267-13274.
  45. Joshi D, Quadri M, Gangane N, Joshi R: **Association of Epstein Barr virus infection (EBV) with breast cancer in rural Indian women.** *PLoS One* 2009, **4**:e8180.
  46. Labrecque LG, Barnes DM, Fentiman IS, Griffin BE: **Epstein-Barr virus in epithelial cell tumors: a breast cancer study.** *Cancer Res* 1995, **55**:39-45.
  47. Xue SA, Lampert IA, Haldane JS, Bridger JE, Griffin BE: **Epstein-Barr virus gene expression in human breast cancer: protagonist or passenger?** *Br J Cancer* 2003, **89**:113-119.
  48. Lake A, Shield LA, Cordano P, Chui DT, Osborne J, Crae S, Wilson KS, Tosi S, Knight SJ, Gesk S, *et al*: **Mutations of NFKBIA, encoding I $\kappa$ B $\alpha$ , are a recurrent finding in classical Hodgkin lymphoma but are not a unifying feature of non-EBV-associated cases.** *Int J Cancer* 2009, **125**:1334-1342.
  49. Osborne J, Lake A, Alexander FE, Taylor GM, Jarrett RF: **Germline mutations and polymorphisms in the NFKBIA gene in Hodgkin lymphoma.** *Int J Cancer* 2005, **116**:646-651.
  50. Levine AJ, Oren M: **The first 30 years of p53: growing ever more complex.** *Nat Rev Cancer* 2009, **9**:749-758.
  51. Oren M, Coors T: **NFKB and p53: A life and death affair.** *Cell Cycle* 2010, **9**.
  52. Oren M, Rotter V: **Mutant p53 gain-of-function in cancer.** *Cold Spring Harb Perspect Biol* 2010, **2**:a001107.
  53. zur Hausen H, de Villiers EM: **Virus target cell conditioning model to explain some epidemiologic characteristics of childhood leukemias and lymphomas.** *Int J Cancer* 2005, **115**:1-5.
  54. Hjalgrim H, Rostgaard K, Johnson PC, Lake A, Shield L, Little AM, Ekstrom-Smedby K, Adami HO, Glimelius B, Hamilton-Dutoit S, *et al*: **HLA-A alleles and infectious mononucleosis suggest a critical role for cytotoxic T-cell response in EBV-related Hodgkin lymphoma.** *Proc Natl Acad Sci USA* 2010, **107**:6400-6405.
  55. Tedeschi R, Bloigu A, Ogmundsdottir HM, Marus A, Dillner J, dePaoli P, Gudnadottir M, Koskela P, Pukkala E, Lehtinen T, Lehtinen M: **Activation of maternal Epstein-Barr virus infection and risk of acute leukemia in the offspring.** *Am J Epidemiol* 2007, **165**:134-137.
  56. Tedeschi R, Pin E, Martorelli D, Bidoli E, Marus A, Pratesi C, Bortolin MT, Zanussi S, Vaccher E, Dolcetti R, De Paoli P: **Serum antibody response to lytic and latent Epstein-Barr virus antigens in undifferentiated nasopharyngeal carcinoma patients from an area of nonendemicity.** *Clin Vaccine Immunol* 2007, **14**:435-441.
  57. Hadinoto V, Shapiro M, Sun CC, Thorley-Lawson DA: **The dynamics of EBV shedding implicate a central role for epithelial cells in amplifying viral output.** *PLoS Pathog* 2009, **5**:e1000496.

doi:10.1186/1743-422X-8-422

**Cite this article as:** Dreyfus *et al*: Analysis of an ankyrin-like region in Epstein Barr Virus encoded (EBV) BZLF-1 (ZEBRA) protein: implications for interactions with NF- $\kappa$ B and p53. *Virology Journal* 2011 **8**:422.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

