A poxvirus protein forms a complex with left-handed Z-DNA: Crystal structure of a Yatapoxvirus Zα bound to DNA


*Department of Molecular Cell Biology, Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon 440-746, Korea; †Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and ‡Department of Biochemistry, College of Medicine, Chung-Ang University, Seoul 156-756, Korea

Contributed by Alexander Rich, July 30, 2004

A conserved feature of poxviruses is a protein, well characterized as E3L in vaccinia virus, that confers IFN resistance on the virus. This protein comprises two domains, an N-terminal Z-DNA-binding protein domain (Zα) and a C-terminal double-stranded RNA-binding domain. Both are required for pathogenicity of vaccinia virus in mice infected by intracranial injection. Here, we describe the crystal structure of the Zα domain from the E3L-like protein of Yaba-like disease virus, a Yatapoxvirus, in a complex with Z-DNA, solved at a 2.0-Å resolution. The DNA contacting surface of Yaba-like disease virus ZαE3L closely resembles that of other structurally defined members of the Zα family, although some variability exists in the β-hairpin region. In contrast to the Z-DNA-contacting surface, the nonbinding surface of members of the Zα family are unrelated; this surface may effect protein-specific interactions. The presence of the conserved and tailored Z-DNA-binding surface, which interacts specifically with the zigzag backbone and syn base diagnostic of the Z-form, reinforces the importance to poxvirus infection of the ability of this protein to recognize the Z-conformation.

The Yaba monkey tumor virus (YMTV) and the closely related Yaba-like disease virus (YLDV) are members of the Yatapoxvirus family. Both viruses have double-stranded DNA genomes of ~150 kilobase pairs and code for >150 proteins (1). YMTV induces histiocytomas with characteristic microscopic morphology in soft tissue of monkeys and baboons (2, 3) and humans (4). YLDV, as well as a third Yatapoxivirus, Tanapox, causes vesicular skin lesions in monkeys and humans (5, 6). Furthermore, YMTV DNA has transforming activity in monkey cell lines (7). The high similarity between the YLDV and YMTV was revealed with the recent sequencing of the Yaba monkey disease virus (8).

All sequenced poxviruses have a gene, called E3L in vaccinia, that is required for IFN resistance. E3L has been shown to be oncogenic and antiapoptotic; NIH 3T3 cells expressing vaccinia E3L were found to grow faster than control cells, with increased expression of cyclin A and decreased levels of the suppressor molecule p26 (9). When the NIH 3T3 E3L-expressing cells were infected by intracranial injection, the ability to bind Z-DNA is essential for E3L activity; domain swaps with the Zα domains from the human ADAR1 and mouse DLM-1 are equivalent to the wild-type, whereas a replacement of the N-terminal domain of E3L with a domain defective in Z-DNA binding results in a less pathogenic or nonpathogenic virus (16). The structures of human ZαADAR1 and mouse ZαDLM1 have been solved, complexed with Z-DNA (11, 17). These structures show that these proteins, which have limited similarity in amino acid sequences, share closely related DNA-binding interfaces. Recently, the NMR structure of the Zα domain of vaccinia E3L (ννZαE3L), solved in the absence of DNA, showed that this viral Zα domain has a similar topology to other Zα domains (18). The E3L orthologue from YLDV has 39% identity to vaccinia virus E3L, spread throughout both N- and C-terminal domains. Although the sequence identity is <30% when only the Zα motifs are compared, all key DNA-contacting residues are conserved. The Zα motifs of YLDV and YMTV are 70% identical (1, 8), reinforcing the close relationship between these two viruses.

Here, we show that the Zα domain of the E3L orthologue of YLDV (yabZαE3L) binds to Z-DNA tightly and specifically in vitro. The crystal structure of a complex between yabZαE3L and Z-DNA is reported. This structure clearly shows that yabZαE3L is a member of the Zα family of protein domains. The interactions between yabZαE3L and Z-DNA are highly conserved with those between ZαADAR1 or ZαDLM1 and Z-DNA. Subtle modulations in the interaction surface demonstrate that variations are allowed in this interaction, particularly at the edge of the binding surface. In contrast, the surface facing away from the Z-DNA interface has no resemblance to other Zα domains. This structure demonstrates that viral Zα domains can bind to the Z-conformation in as well engineered a fashion as previously characterized Zα domains from mammalian proteins.

Materials and Methods

Expression and Purification. The gene coding for the Zα domain (residues 2–79) of E3L from YLDV (yabZαE3L) was subcloned into pET28a (Novagen) and transformed into BL21(DE3) (Novagen). The cells were grown in LB containing 30 µg/ml kanamycin to an OD600 of 0.7–0.8 at 30°C, and then induced with 0.6 mM isopropyl β-D-thiogalactoside (IPTG) at 18°C for 20 h. yabZαE3L was purified by using a slight modification of the

Abbreviations: Zα, Z-DNA-binding protein domain; yabZαE3L, Zα domain from the E3L-like protein of Yaba-like disease virus; YMTV, Yaba monkey tumor virus; YLDV, Yaba-like disease virus.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1SU7). To whom correspondence may be addressed. Fax: 617-253-8699 (A.R.). E-mail: kkim@med.sskku.ac.kr (K.K.K.) or ygkim@cau.ac.kr (Y.-G.K.).

© 2004 by The National Academy of Sciences of the USA
method described for purifying hZADAR1 (17). After a His-affinity column (Amersham Pharmacia Biosciences) and removal of the N-terminal six histidines with thrombin (Boehringer Mannheim), the protein in buffer A (20 mM Hepes, pH 7.5/80 mM NaCl) was further purified by using Resource S (Amersham Pharmacia Biosciences). The purified protein was dialyzed against buffer A and concentrated to 1 mM by using YM-10 ultrafiltration and Centricon (Millipore). All of the steps of buffer exchange were performed by using a PD-10 column (Amersham Pharmacia Biosciences). The protein concentration was measured by using the Bradford method with a BSA standard curve (19).

CD.

(dC-dG)₆ (IDT) was rehydrated with a TE buffer (10 mM TrisCl, pH 8.0/1 mM EDTA) and annealed before use. The conversion of (dC-dG)₆ to Z-DNA was monitored by CD. The CD spectra were obtained at 25°C by using an AVIV model 202. All measurements were carried out by using 50 μM/ml (75 μM base pair) DNA in a CD buffer (10 mM Hepes, pH 7.4/10 mM KF/0.1 mM EDTA) in a 2-mm quartz cell. Protein was added to the DNA solution to a final concentration of 30 μM. The maximum volume of the protein added to the sample did not exceed 5% of the total volume. Spectra for wavelength scanning were recorded at 1-nm intervals, averaged over 3 s. For the kinetic measurements, the CD signal changes at 255 nm were recorded at 0.5-s intervals up to 1 h.

Crystallization and Data Collection.
The Zα domain of YLDV and d(TCGCGCG)₂ (IDT, Coralville, IA) were purified, mixed and crystallized (S.C.H., D.B.O., Y.-G.K., and K.K.K., unpublished data). Briefly, the protein and DNA in 5 mM Hepes-NaOH (pH 7.5) and 80 mM NaCl were mixed at an equimolar (protein:single strand DNA) ratio, resulting in a final concentration of 0.6 mM each, and incubated at 30°C for at least 10 h. Initial crystallization was performed at 22°C in hanging drops by using Hampton screens I and II and Natrix screen kits, and the vapor diffusion method (Hampton Research, Riverside, CA). The protein–DNA complex (1.5 μl) was mixed with 1.5 μl of a reservoir solution containing 25 mM cacodylate-HCl (pH 6.0), 1–1.1 M LiSO₄.
and 2–4 mM MgCl2 after 1 month. Br-multinwavelength anomalous dispersion (MAD) x-ray data were collected at the 41XU beam line of the Spring-8 Synchrotron by using a crystal soaked in a solution of 25 mM cacodylate-HCl (pH 6.0), 1 M NaBr, and 25% glycerol for 50 s. All data were processed by using the HKL2000 program (20). Crystals belong to the orthorhombic space group P21212, with unit cell constants a = 51.20, b = 92.45, and c = 48.02 Å (Table 1, which is published as supporting information on the PNAS web site).

Structure Determination and Refinement. The Br sites were found, and the initial multinwavelength anomalous dispersion (MAD) phases were calculated and refined by using the programs SOLVE and RESOLVE, respectively (21, 22). The MAD phases were further refined by using solvent flattening, histogram matching, and non-crystallographic symmetry averaging by using the program DM (23).

The quality of the final map was sufficient to fit into the model. The model building was performed with the O program (24). The initial model of the protein and DNA was built by locating the Zα and DNA of hADAR1 (chain C and F; PDB code 1OBJ) in a final electron density map. The model was refined with the data set collected at 2.0 Å from the same crystal used for the MAD data collection by using the CNS program (version 1.0) with the MLHL target function. The model was then rebuilt against an initial map, a phase combined 2F – F and an F – F map (25). The stereochemistry of the model was checked with the program PROCHECK at each stage of model building and refinement (26). There were two proteins (chains A and B) and one double-stranded DNA (chains C and D) in the asymmetric unit. Alanine and glycine residues were fitted to disordered regions with poor electron density (Ala-7, Ala-8, and Ala-9 in chain A; Gly-40 in chain B). The final model included the residues, 6–75 of each protein, 12 nucleotides, and 171 waters (Table 1).

Results and Discussion

Binding of yabZE3L to Z-DNA. There are nine residues that make specific contacts with Z-DNA in the co-crystal structure of hZADAR1:DNA. Eight of these residues are conserved in yabZE3L (Fig. 1A). Therefore, we thought it likely that yabZE3L would bind to Z-DNA tightly and specifically. This binding was tested by using CD (Fig. 2). The CD spectrum in the range between 240 and 300 nm of Z-DNA is almost a mirror image as compared with that of B-DNA (27). When (dC-dG)6 is incubated with human ZADAR1 or mouse ZADLM1, the spectrum of the DNA rapidly converts to that of Z-DNA (10, 12, 28). In the presence of yabZE3L, a similar conversion takes place (Fig. 2A). The change from B- to Z-DNA over time can be monitored at 255 nm, as shown in Fig. 2B. The yabZE3L domain induces the conversion of B- to Z-DNA as quickly as that of human ZADAR1, whereas a somewhat slower conversion was observed with the mouse ZADLM1 (Fig. 2B). A comparative study of binding affinities will be presented elsewhere. These spectral results show that the yabZE3L domain stabilizes the Z-conformation in a manner similar to that seen with the other two domains. Although yabZE3L and human ZADAR1 share only 26% sequence identity (Fig. 1A), Z-DNA binding by both reinforces the idea that the major factor determining Z-DNA-binding ability mainly resides in the Z-DNA contacting amino acid residues.

Overall Structure of the yabZE3L-Z-DNA Complex. To further investigate the nature of the interaction between yabZE3L and DNA, the three-dimensional structure of co-crystals of yabZE3L and the oligonucleotide, d(TCGCGCG), was determined to 2.0 Å with an R factor of 23.7% and an Rfree of 27.0% (Table 1). In the crystal, two yabZE3L domains are found in the asymmetric unit, each bound to a strand of double-stranded DNA in the Z-conformation (Fig. 1B). The proteins do not interact with each other. One Zα domain and its interacting DNA were designated as chains A and C, the other protein–DNA pair were designated as chains B and chain D. The two yabZE3L domains are related by noncrystallographic symmetry and are nearly identical, with 0.47 Å of rms deviation (rmsd) for 64 Cα atoms. The crystal structure of yabZE3L clearly verifies that it belongs to the Zα family of helix–turn–helix (HTH), winged-helix Z-DNA-binding proteins. It has an α/β architecture, consisting of three β-strands and three α-helices, as is found in hZADAR1 and mZADLM1 (Fig. 1B). The three α-helices (α1, α2, and α3) form a core domain, which is flanked by a β-sheet of three antiparallel strands (β1, β2, and β3). α2 and α3 form an HTH motif and two antiparallel β-strands (β2 and β3) form the wing.

In the complex, the double-stranded DNA adopts a typical left-handed Z-DNA conformation, with guanine nucleotides in the syn conformation and C3-endon sugar puckering except G6, which is in C2-endon. Watson–Crick base pairing is well conserved in this structure. The S’-T overhang is not modeled due to its weak electron density.

Protein–DNA Interaction in yabZE3L. The interaction between members of the Zα family and Z-DNA is characterized by a single continuous recognition surface involving numerous sugar-phosphate backbone contacts, which provide conformational specificity (11, 17). This recognition surface is made up of residues in the
α3 helix and the "wing." DNA recognition by yabZ_E3L has these same properties. Three residues, Asn-47, Tyr-51, and Trp-69 as numbered in yabZ_E3L (Asn-173, Tyr-177, and Trp-195 in hZ_ADAR1), central to interaction with Z-DNA, are completely conserved within the Zα family (Figs. 1 and 3). Asn-47 and Trp-69 make water-mediated hydrogen bonds to the phosphate backbone. Trp-69 also buttresses Tyr-51 by using a distinctive hydrophobic edge-to-face interaction, thereby stabilizing the interaction of Tyr-51 with the C8 of the syn guanine G4. In addition to this CH-π interaction, Tyr-51 makes a direct hydrogen bond to a phosphate in the DNA backbone. Glu-48, Lys-43, and Lys-44 from the recognition helix of helix–turn–helix motif also participate in DNA recognition, either directly or through water-mediated hydrogen bonds to the phosphate backbone (Fig. 3).

The importance of the interaction of tyrosine on the recognition helix with DNA has been demonstrated in a biological system assessing vaccinia virus infection of mice. Wild-type virus is lethal for the mouse in intracranial injection. When hZ_ADAR1 is substituted for vvZ_E3L, the chimeric virus retains lethality. However, mutating the recognition helix (Y to F, Y to A) results in significant loss of lethality for both the wild-type and the hZ_ADAR1 chimeric virus, as well as weaker Z-DNA binding in vitro (16). The Tyr-51 of yabZ_E3L shows a conformation similar to that found in the hZ_ADAR1:Z-DNA complex.

In this crystal, the asymmetric unit contains two yabZ_E3L and one double-stranded DNA. Although the contacts between yabZ_E3L and each strand of DNA use the same residues and modes, the details of the interactions are slightly different. The most noticeable difference is found in the interaction between the wing hairpin and DNA. In this region, Pro-66, Pro-67, and Asn-65 mediate interactions between Zα and Z-DNA; however, the Asn-65-DNA contact differs between the two forms. In one binding surface (chains B and D), Asn-65 interacts with the phosphate and ribose directly, whereas, in the other side (chains A and C), the hydrogen bonds are water-mediated (Fig. 3). This finding indicates some flexibility in the β-sheet–DNA contacts. The position of the β-turn in the solution structure of vvZ_E3L is considerably different from that of the yabZ_E3L wing (Fig. 1C) (18). It is possible that, in viral proteins, the wing moves as it clasps the Z-DNA and that the two different interactions seen in this crystal are the result of this movement. On the other hand, the position and sequence of the wing in mZ_DLM1, which has only one proline in the β turn that contacts Z-DNA, is also different from that of hZ_ADAR1 and yabZ_E3L. The position of the vaccinia wing may be a further example of the variability of this region.

Structural Comparison of yabZ_E3L:Z-DNA with other Zα:Z-DNA Complexes. The overall structure of yabZ_E3L and its interactions with Z-DNA are very similar with those of hZ_ADAR1 and mZ_DLM1.
addition, the protein:DNA contacts are similar to those seen in the hZαADAR1, mZαDLM1, and yabZαE3L :Z-DNA structures (18). The hZαADAR1:Z-DNA complex can be superposed on the yabZαE3L :Z-DNA complex with an rms deviation (rmsd) of 1.10 Å for 64 Cα atoms and six nucleotides, whereas the rmsd between the yabZαE3L :Z-DNA complex and the mZαDLM1:Z-DNA complex is 1.37 Å when 57 Cα atoms and six nucleotides were used for calculation. Therefore, yabZαE3L is more similar to hZαADAR1 than to mZαDLM1. The most prominent differences between these structures are seen in the wing (β2 and β3). mZαDLM1 has a shorter wing, which is positioned closer to the α3 helix, whereas the wings of yabZαE3L and hZαADAR1 in complex with Z-DNA are nearly identical (Fig. 1C). It is possible that the slower B- to Z- conversion rate observed for mZαDLM1 is related to the different shape of the wing (Figs. 1C and 2B).

In vvZαE3L, the solution structure shows that Tyr 48, orthologous to Tyr-51 of yabZαE3L, adopts a different side chain conformation in the absence of Z-DNA. However, when vvZαE3L is bound to DNA, the tyrosine adopts a configuration similar to that seen in hZαADAR1, suggesting that the Z-DNA contacting surface of vvZαE3L is very similar when the protein is bound to Z-DNA. The requirement for a change in conformation of the protein upon binding may result in the slower binding kinetics of vvZαE3L as compared to other Zα domains.

When bound to DNA, a surface area of 744 Å² is buried in yabZαE3L, comparable with 725 Å² for hZαADAR1 and 740 Å² for mZαDLM1 (11), which suggests that these three Zα proteins should have similar binding affinities for Z-DNA.

The Z-DNA-binding surfaces of the four Zα domains whose three-dimensional structure has been determined show very similar curvature and charge distribution (Fig. 4 Upper), making a tailored fit for the ligand. The first three (yabZαE3L, hZαADAR1, mZαDLM1) were solved bound to Z-DNA, and the bound strand is shown on the surface. The fourth protein, vvZαE3L, was solved in solution without Z-DNA. Analysis of its interactions with Z-DNA shows that the DNA occupies the same position as in the first three diagrams. In each case, the central binding surface is lined with positive charges, especially around the periphery of the binding site. A white “bulge” in the center of the first three figures represents the tyrosine on the recognition helix whose face interacts with a syn guanine residue. That bulge is missing in the vvZαE3L, because the tyrosine has rotated to the top of the binding surface. The position of this tyrosine is variable in solution; Fig. 4 shows the predominant position, in which the tyrosine prevents binding to Z-DNA. Binding to Z-DNA requires the tyrosine to rotate down into the center of the binding surface.

It is interesting that the distribution of positive charges around the binding site in the top row varies somewhat in all four structures. Fig. 4 Lower shows the nonbinding “back” side of the Zα domains, on which an apparently random distribution of positive and negative charges is seen. This result is not surprising because these proteins are likely to have other functions in addition to Z-DNA binding, even though many of these functions are not known at present.

Conclusion

In this paper, we have described the interactions of the Zα domain of the E3L-like protein of YLDV with DNA by using spectroscopic and crystallographic studies. This E3L-like protein is known to play a key role in poxvirus pathogenesis, as is best documented in vaccinia virus infections. The yabZαE3L readily converts B-DNA to Z-DNA and is comparable in activity to Zα domains from hADAR1 and mDLM1 (Fig. 2C). The overall structure of yabZαE3L and its Z-DNA recognition in the complex are extremely similar to those of the mZαDLM1:Z-DNA and hZαADAR1:Z-DNA complexes, because yabZαE3L also uses a winged helix-turn-helix motif to recognize the zigzag backbone and syn base of Z-DNA. The critical and conserved residues of yabZαE3L for Z-DNA recognition also have similar interactions for Z-DNA recognition at the same position. In addition, a few different structural features found among the Zα domains might be related to their different roles in Z-DNA binding and related functions in the cell.

The presence of a protein, precisely tailored to fit Z-DNA, supports the importance in poxviruses of binding to this unusual conformation. Although yabZαE3L resembles previously studied Zα domains in every way, it is also important to notice the differences between these domains, which may be related to specific function.
This work was supported by Korea Research Foundation Grant KRF-2003-015-E00058 (to K.K.K.), a National Institutes of Health grant, and grants from the Dana and Ellison Medical Foundations (to A.R.). The