The solution structure of the N-terminal domain of E3L shows a tyrosine conformation that may explain its reduced affinity to Z-DNA in vitro

Jan D. Kahmann*, Diana A. Wecking*, Vera Putter*, Ky Lowenhaupt†, Yang-Gyun Kim‡, Peter Schmieder§, Hartmut Oschkinat$, Alexander Rich**, and Markus Schade***

*Combibiotics Biopharm AG and **Forschungsinstitut für Molekulare Pharmakologie, Robert-Rösle-Strasse 10, D-13125 Berlin, Germany; †Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139; and §Department of Biochemistry, College of Medicine, Chung-Ang University, 221 Heusuk-Dong, Dongjak-ku, Seoul 156-756, Korea


The N-terminal domain of the vaccinia virus protein E3L (Z\textsubscript{E3L}) is essential for full viral pathogenicity in mice. It has sequence similarity to the high-affinity human Z-DNA-binding domains Z\textsubscript{ADAR1} and Z\textsubscript{DLM1}. Here, we report the solution structure of Z\textsubscript{E3L} and the chemical shift map of its interaction surface with Z-DNA. The global structure and the Z-DNA interaction surface of Z\textsubscript{E3L} are very similar to the high-affinity Z-DNA-binding domains Z\textsubscript{ADAR1} and Z\textsubscript{DLM1}. However, the key Z-DNA contacting residue Y48 of Z\textsubscript{E3L} adopts a different side chain conformation in unbound Z\textsubscript{E3L}, which requires rearrangement for binding to Z-DNA. This difference suggests a molecular basis for the significantly lower in vitro affinity of Z\textsubscript{E3L} to Z-DNA compared with its homologues Z\textsubscript{ADAR1} and Z\textsubscript{DLM1}.

Materials and Methods

Protein Preparation. Residues 1–78 of the vaccinia virus gene E3L (GenBank no. A0A07259), comprising the Z\textsubscript{E3L} domain, with four additional vector-encoded residues at the N terminus, were expressed as a fusion protein with an N-terminal (His)\textsubscript{6} tag, tag from a pET-28 vector (Novagen) in E. coli strain BL21(DE3). Alternatively, residues 1–70 of E3L were expressed similarly for DNA interaction assays.

To produce \textsuperscript{15}N- and \textsuperscript{15}N/\textsuperscript{13}C-labeled Z\textsubscript{E3L}, bacteria were grown in M9 medium containing 1 g/liter \textsuperscript{15}NH\textsubscript{4}Cl and 1.5 g/liter \textsuperscript{13}C-glucose. Cultures were induced with 1 mM isopropyl \textbeta-D-thiogalactoside (IPTG) when they reached OD\textsubscript{600} of 0.8. After 4 h incubation, cells were harvested, resuspended in 20 mM NaH\textsubscript{2}PO\textsubscript{4}, pH 8.0, and dialyzed overnight at room temperature in PBS. The cleaved protein was loaded on a Resource Q column (Amersham Biosciences), and the bound protein was eluted with a gradient of 0–1 M NaCl in 20 mM NaH\textsubscript{2}PO\textsubscript{4}, pH 6.5. Alternatively, bacteria were lysed with Bugbuster (Novagen) following the manufacturer's protocol, and protein was purified as described.

Abbreviations: Z\textsubscript{a}, Z-DNA-binding protein domain; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser enhancement.

Data deposition: The atomic coordinates of the 20 lowest energy structures of a total of 450 have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1OV9).

© 2004 by The National Academy of Sciences of the USA

PNAS | March 2, 2004 | vol. 101 | no. 9

DOI: 10.1073/pnas.0306612100
of flight (MALDI-TOF) mass spectrometry of the 15N- and 13C-labeled Z$_{E3L}$ yielded single peaks at 9,291 Da and 9,672 Da, respectively, which agree very well with the calculated molecular masses of 9,292 Da and 9,685 Da, respectively.

**NMR Spectroscopy.** NMR experiments were carried out at 25°C on 2.2 mM U-15N-labeled and 2 mM U-13C,15N-labeled Z$_{E3L}$ samples in 20 mM sodium-phosphate (pH 6.5), 20 mM NaCl, 0.1 mM Na$_3$N$_5$ with 5% and 100% D$_2$O, respectively, on 600-MHz NMR spectrometers. 1H, 15N, and 13C resonance assignments were obtained from the following 3D heteronuclear correlation experiments (10): CBCA(CO)NH, CBCANH, HBBHA(CO)NH, HBC(OC)'SA, CH(CO)CONH, CH(CO)OSY, and HCCH-TOCSY. Interproton distance restraints were derived from 3D 15N-heteronuclear single quantum coherence (HSQC)-NOESY (150-ms mixing time), 3D 13C-HSQC-NOESY (40 and 70 ms mixing times). Spectra were processed with XWINNMR (Bruker) and analyzed with SPARKY 3.105 (11). Spectra were referenced by external calibration on 2,2-dimethyl-silapentane-5-sulfonic acid (DSS), sodium salt (12).

**Interaction Mapping.** For interaction mapping, a shortened Z$_{E3L}$ construct (comprising residues 5–70 of GenBank no. AAA02759) was used that lacks the first four N-terminal and the last eight C-terminal residues. These residues are nonstructured in the 3D structure of Z$_{E3L}$. The 1H and 15N backbone chemical shifts are virtually identical between this construct and the 1–78 residues construct, indicating that both constructs share the same 3D fold. 1D 1H and 2D 15N-HSQC NMR spectra were recorded on the following four samples in 20 mM [bis(2-hydroxyethyl)amino]tris(hydroxymethyl)methane (Bistris; pH 6.7), 50 mM NaCl, 5% D$_2$O at 25°C: (i) 40 mM Z$_{E3L}$; (ii) 40 mM Z$_{E3L}$ with 300 mM [Co(NH$_3$)$_6$]Cl$_2$; (iii) 40 mM Z$_{E3L}$ with 10 mM d(CG)$_5$T$_4$(CG)$_6$; and (iv) 40 mM Z$_{E3L}$ with 10 mM d(CG)$_5$T$_4$(CG)$_6$ and 300 mM [Co(NH$_3$)$_6$]Cl$_2$. After NMR data acquisition, CD spectra were recorded on each sample at room temperature by using a 2-mm cuvette on a Jasco J-720 CD spectrometer (Jasco, Easton, MD). CD control spectra of 10 mM d(CG)$_5$T$_4$(CG)$_6$ in the B-DNA conformation and in the Z-DNA conformation were recorded in a 1-mm cuvette at room temperature in 20 mM Bistris (pH 6.7), 50 mM NaCl buffer alone, and with 5 M NaCl, respectively. The assignments and concentration dependence of the chemical shift changes of Z$_{E3L}$ were confirmed by a titration experiment on a 20-µM Z$_{E3L}$ sample with increasing concentrations of d(CG)$_5$T$_4$(CG)$_6$ (1, 2, 4, 7, and 12 µM) and a constant [Co(NH$_3$)$_6$]Cl$_2$ over DNA excess of 30 in 20 mM Bistris (pH 6.7), 50 mM NaCl, and 5% D$_2$O at 25°C. Spectra were processed and analyzed as described above. Chemical shift changes were averaged according to the formula \[ ((\Delta\text{H})^2 + (\Delta\text{N}/5)^2)^{1/2} \] (13).

**Structure Calculation.** Nuclear Overhauser enhancement (NOE) distance restraints derived from 15N- and 13C-edited NOE-SY experiments were manually assigned and further analyzed (calibration and removal of redundant distance restraints) by using the program DYANA 3.1 (14). Seventy-two backbone dihedral angle constraints were derived from Ca chemical shifts according to the rules (15): $-120^\circ < \psi < -20^\circ$ and $-100^\circ < \phi < 0^\circ$ for $\Delta(Ca) > 1.5$ ppm, and $-200^\circ < \psi < -80^\circ$ and $40^\circ < \phi < 220^\circ$ for $\Delta(Ca) < -1.5$ ppm. The dihedral angle constraints are in agreement with the preliminary structure calculated solely from the NOE restraints. Further 20 hydrogen bonds within α-helices and four hydrogen bonds within β-strands were derived from the NOE-based preliminary structure and confirmed by the analysis of the Ca and CB chemical shift values by using the program TALOS (16). Structures were calculated by 4,000 steps of simulated annealing with torsion angle dynamics and subsequently 1,000 steps of minimization in DYANA 3.1. For better convergence during structural refinement, the ψ and φ dihedral angles of the residues 59, 60, and 61 preceding the cis peptide bond between I62 and P63 were preset to a range that is wider by 5° or more than the range of the structural ensemble calculated without such preset angles.

**Results and Discussion**

**Structure Determination.** The 3D structure of the N-terminal Z-DNA-binding domain of the vaccinia virus gene product E3L (residues 1–78) was determined by multidimensional NMR spectroscopy in solution. Complete chemical shift assignments were obtained from 3D triple-resonance and double-resonance NMR spectra except for the first two vector-encoded residues. Chemical shifts of residues 1–10 and 69–78 are not well dispersed. This finding is confirmed by the 3D structure demonstrating that the folded core domain comprises residues 11–68, whereas all other N- and C-terminal residues are unstructured. Structural statistics are listed in Table 1. The coordinates of the ensemble of the 20 lowest energy structures of Z$_{E3L}$ (residues 9–70) have been deposited in the Protein Data Bank with the PDB ID code 1OYI.

**Structure Description.** The Z$_{E3L}$ domain is composed of three α-helices (designated α1, α2, and α3) and three β-strands (designated β1, β2, and β3) in an α1β1α2α3β2β3 linear order (Fig. 1). Helices α1 and α3 pack against a short anti-parallel, triple-stranded β-sheet, in which β3 is sandwiched between β1 and β2. Strands β1 and β3 are connected by only two backbone hydrogen bonds between residues A27 and W66. Strands β2 and β3 are bridged by three hydrogen bonds comprising residues Y57 and S59 of β2 and R65 and F67 of β3. The ensemble of the 20 lowest energy structures shows that only loop 2 between α2 and α3 is less rigid whereas all other loops are tightly structured so that Z$_{E3L}$ is a rigid body. Loop 4 between β2 and β3 in α3 is made rigid by the two sequential prolines 63 and 64, of which the former adopts a rare cis peptide bond. The side chains of the
other residues in this loop (D60, D61, I62, and R65) are flexible in the structural ensemble, giving this loop the shape of a solvent accessible finger with a rigid backbone. Mutational (5) and structural studies (7, 9) have demonstrated that this distinctly conserved feature is essential for selective interaction of the homologous Z\textsubscript{ADAR1} domain with Z-DNA. In the co-crystal structure of Z\textsubscript{ADAR1} and Z-DNA, the protein makes several important van der Waals contacts to Z-DNA.

**Chemical Shift Mapping of the E3L/Z-DNA Interaction Surface.** Alternating d(CG)\textsubscript{n} oligomers have been successfully used to study the interaction of the Z-DNA-binding domains Z\textsubscript{ADAR1} (7, 9) and Z\textsubscript{DLM1} (8) with Z-DNA. In contrast to these high-affinity Z-DNA-binding homologues, Z\textsubscript{E3L} does not flip the B-DNA conformation of these substrates into the Z-conformation when incubated with each other under physiological buffer conditions at micromolar concentrations (3). The CD spectrum of 40 \mu M Z\textsubscript{E3L} in the presence of 10 \mu M d(CG)\textsubscript{n}T\textsubscript{d}(CG)\textsubscript{n} clearly indicates a B-DNA conformation for this DNA substrate (Fig. 2A, blue curve). Further, the 2D \textsuperscript{15}N-HSQC NMR spectrum of this sample shows no chemical shift changes when compared with the spectrum of the protein alone (see Fig. 4, which is published as supporting information on the PNAS web site). This result indicates that Z\textsubscript{E3L} does not interact with d(CG)\textsubscript{n}T\textsubscript{d}(CG)\textsubscript{n} in the B-DNA conformation under these conditions because chemical shifts are sensitive even to weak interactions.

To investigate the interaction between Z\textsubscript{E3L} and d(CG)\textsubscript{n}T\textsubscript{d}(CG)\textsubscript{n} in the Z-DNA conformation, [Co(NH\textsubscript{3})\textsubscript{6}]\textsuperscript{3+}, which is known to promote the conversion from B- to Z-DNA (17), was added to a final concentration of 300 \mu M under otherwise identical conditions. The CD spectrum of this sample shows large alterations of the molar ellipticity at 295 and 255 nm, which are characteristic for d(CG)\textsubscript{n} oligomers in a left-handed Z-DNA conformation (Fig. 2A, red curve). The corresponding \textsuperscript{15}N-HSQC NMR spectrum shows a large number of chemical

---

**Fig. 1.** 3D solution structure of Z\textsubscript{E3L}. (A) Stereoview of the ensemble of the 20 lowest energy structures of E3L. The first and last residue of the 3 \alpha-helices (red) (\alpha\textsubscript{1}, \alpha\textsubscript{2}, \alpha\textsubscript{3}) and 3 \beta-strands (cyan) (\beta\textsubscript{1}, \beta\textsubscript{2}, \beta\textsubscript{3}) are numbered. (B) Stereoview of the backbone ribbon of the mean structure illustrating the (\alpha plus \beta) helix-turn-helix fold of Z\textsubscript{E3L}. The N- and C-termini are labeled with N and C, respectively. (C) The secondary structure of Z\textsubscript{E3L} is shown with the amino acid sequence directly under it.
shift changes and three vanishing signals (Fig. 2B) with respect to ZOEEL alone, indicating that ZOEEL interacts with d(CG)_6T_4(CG)_6 in the Z-DNA conformation. The concentration dependence of these chemical shift alterations has been confirmed by an independent chemical shift-mapping experiment at 20 μM ZOEEL with increasing d(CG)_6T_4(CG)_6 concentration.
the loops connecting \( \text{H}9251 \)/21 and three structures of \( \text{E3L} \) (blue), unbound \( \text{Z} \) (red), and bound \( \text{Z} \) (light green). Whereas \( \text{Y}48 \) residues are within van-der-Waals distance to \( \text{W}66 \) in both \( \text{Z} \) structures, it adopts two solvent exposed rotamer positions in \( \text{Z} \), which are 7.2 and 10.8 Å apart from \( \text{W}66 \).

The Backbone Structure of \( \text{Z} \) Is Similar to \( \text{Z} \) and \( \text{Z} \), As expected from the primary sequence homology between \( \text{Z} \), \( \text{Z} \), and \( \text{Z} \), the three Z-DNA-binding domains share the same \( \alpha/21\beta/223/23 \) topology. The structure of \( \text{Z} \) shows a backbone rms deviation of 1.24 Å to \( \text{Z} \) and of 1.21 Å to \( \text{Z} \) (superposition of helices and strands only), indicating that the sequence homology is paralleled by a high overall structural homology. In particular, the three \( \alpha/21 \)-helices and three \( \beta/223 \)-strands overlay very well between \( \text{Z} \), \( \text{Z} \), and \( \text{Z} \) (Fig. 3C). Structural differences are observed for the loops connecting \( \alpha/1 \) and \( \beta/1 \) (loop 1), \( \alpha/2 \) and \( \alpha/3 \) (loop 2), and \( \beta/2 \) and \( \beta/3 \) (loop 4), of which the latter shows the most marked deviation in its backbone conformation. This finding is not unexpected because loop 4 contains profound differences on the primary sequence level. In \( \text{Z} \) loop 4 (all subsequent amino acid numbers refer to the homologous residues of \( \text{Z} \)) is shorter by two residues than in \( \text{Z} \) and \( \text{Z} \). Moreover, the six residues preceding \( \text{P}63 \) of loop 4 are poorly conserved between \( \text{Z} \) and \( \text{Z} \). The cis proline 63 of this loop is the sole conserved residue in \( \text{Z} \), \( \text{Z} \), and \( \text{Z} \). Proline 63 is of particular importance for the Z-DNA-binding activity because it confers direct Z-DNA contacts in the co-crystal structures of \( \text{Z} \) (7) and \( \text{Z} \) (8). Furthermore, the strongest loss of virulence is found when this residue is mutated to alanine in wild-type \( \text{Z} \) (4). In the 3D structures of \( \text{Z} \) and \( \text{Z} \), \( \text{P}63 \) adopts identical positions at the tip of loop 4 although in \( \text{Z} \) the entire loop 4 is rotated away from helix \( \alpha/3 \), resulting in a distance of 5.9 Å between the N atoms of \( \text{P}63 \) of \( \text{Z} \) and \( \text{Z} \). This offset in the interaction surface may be compensated by a subtle adjustment in the binding geometry between \( \text{Z} \) and Z-DNA. In conclusion, the overall backbone structure of \( \text{Z} \) is very similar to its homologues \( \text{Z} \) and \( \text{Z} \), with the exception of loop 4, which shows a displacement that is not expected to markedly affect binding to Z-DNA.

Y48 Adopts a Distinct Conformation in \( \text{Z} \). To provide a molecular understanding for the significantly lower affinity of \( \text{Z} \) to Z-DNA as compared with \( \text{Z} \) and \( \text{Z} \), the structural comparison of the Z-DNA contacting residues between these homologues is of paramount interest. Of the total of eight common Z-DNA-contacting residues in the co-crystal structures of \( \text{Z} \) and \( \text{Z} \), the three residues \( \text{P}63 \), \( \text{P}64 \), and \( \text{W}66 \) possess rigid side chains whose conformations are identical between \( \text{Z} \) and \( \text{Z} \) within the limitations of the backbone comparison described above. Of the remaining five Z-DNA contacting residues in helix \( \alpha/3 \), the side chains of \( \text{K}40 \), \( \text{R}41 \), \( \text{N}44 \), and \( \text{K}45 \) adopt similar positions in \( \text{Z} \), \( \text{Z} \), and \( \text{Z} \) (Fig. 34). Only the side chain of \( \text{Y}48 \) is markedly different, showing a distinct solvent exposed conformation in \( \text{Z} \). A closer view of \( \text{Y}48 \) and \( \text{W}66 \) in the ensemble of the 20 lowest energy structures of \( \text{Z} \) superimposed on \( \text{Z} \) and \( \text{Z} \) demonstrates (Fig. 3B) that the phenolic ring of \( \text{Y}48 \) probably adopts two major rotamer positions, which are 7.2 Å and 10.8 Å apart from \( \text{W}66 \) (distance \( \text{Y}48(C2) - \text{W}66(C2) \)). In
contrast, the Y48–W66 distance in bound (9) and unbound \( Z_{\text{ADAR1}} \) (7) measures only 3.9 and 4.5 Å, respectively. The experimental foundation of this marked difference is the observation of several long-range NOEs between the aromatic rings of Y48 and W66 in the NMR structure of unbound \( Z_{\text{ADAR1}} \) but none of such NOEs in the \( ^{13} \text{C} \)-edited NOESY spectra of \( Z_{\text{ADHEL}} \). Neither does Y48 of \( Z_{\text{ADHEL}} \) show long-range NOEs to other residues. The general sensitivity of the NOESY experiments on \( Z_{\text{ADHEL}} \) is confirmed by the observation of 21 long-range NOEs to the aromatic protons of W66. Therefore, Y48 of \( Z_{\text{ADHEL}} \) adopts two flexible solvent-exposed rotamer positions whereas Y48 of \( Z_{\text{ADAR1}} \) is tightly packed against W66 in both the unbound and bound state.

In the co-crystal structures of \( Z_{\text{ADAR1}} \) and \( Z_{\text{ADHEL}} \), Y48 is the only residue that mediates direct contacts to a base of the bound Z-DNA. In this interaction, the base adopts the syn conformation, which is characteristic for the left-handed Z-conformation of double-stranded DNA. This close interaction geometry suggests that in \( Z_{\text{ADHEL}} \), the solvent-exposed side chain of Y48 rearranges when \( Z_{\text{ADHEL}} \) binds to Z-DNA. By analyzing resolved \( ^{1} \text{H} \) chemical shifts in 1D\(^{1} \text{H} \)-NMR spectra of 20 μM \( Z_{\text{ADHEL}} \) with increasing concentrations of \( d(\text{CG})_5\text{T}(\text{CG})/ [\text{Co(NH}_3)_6]^{3+} \), we found that the aromatic H6 and H8 atoms of Y48 vanish when \( Z_{\text{ADHEL}} \) binds to Z-DNA. Moreover, the chemical shifts of both methyl groups of L47 (−0.21 and −0.314 ppm in unbound \( Z_{\text{ADHEL}} \)), which are packed in van-der-Waals distance underneath the indole ring of W66, alter in this experiment. These data indicate that the chemical environment around the side chains of L47, Y48, and W66 changes when Z-DNA is bound. Further, the observation of selectively vanishing signals for the HN, H6, and H8 of Y48 and the HN of K45 (Fig. 5), which is connected to HN of Y48 through an \((i, i + 3)\) α-helical hydrogen bridge, is in agreement with conformational rearrangements of the Y48 when \( Z_{\text{ADHEL}} \) binds to Z-DNA. The cost in energy for such a rearrangement may account for the substantially lower affinity of \( Z_{\text{ADHEL}} \) to Z-DNA, as compared with \( Z_{\text{ADAR1}} \) and \( Z_{\text{ADHEL}} \), where Y48 is prepositioned to bind Z-DNA (9).

Mutational experiments suggest that Y48 plays a key role for both binding to Z-DNA in vitro as well as viral pathogenicity in mice. In \( Z_{\text{ADAR1}} \), the mutation of Y48 to alanine leads to both a profound loss in Z-DNA affinity and a significant reduction in binding specificity to the Z-conformation of DNA, as evidenced by Bicore and CD spectroscopy (4, 5, 7). In \( Z_{\text{ADHEL}} \), the mutation of Y48 to alanine abrogates viral pathogenicity in mice by three \( \log_{10} \) units (4). It is therefore intriguing to consider Y48 a conformational switch that has to be turned inward toward the protein to enable binding to Z-DNA. In vivo, this switch may be turned on by activating proteins and induced by preformed segments of Z-DNA.

The importance of the tyrosine–Z-DNA interaction is further illustrated by a second domain in ADAR-1 called \( Z_{\text{BADAR1}} \). Although it has many sequence similarities to \( Z_{\text{ADAR1}} \), it lacks the tyrosine on helix 3, has an isoenceule instead, and shows no in vitro Z-DNA binding (3, 4). When put into vaccinia virus instead of \( Z_{\text{ADHEL}} \), the chimeric virus shows no pathogenicity (4). However, if a mutant \( Z_{\text{BADAR}} \) is made, changing isoenceule to tyrosine, it then binds Z-DNA in vitro, and the chimeric virus becomes pathogenic.

A yeast one-hybrid system has been developed in which reporter gene (β-galactosidase) expression depends on binding of a protein to Z-DNA near the promoter (3). The protein is fused to a transcriptional activator domain, which turns on the gene. When \( Z_{\text{ADHEL}} \) is used, the response of the reporter gene is the same as when \( Z_{\text{ADAR1}} \) or \( Z_{\text{ADHEL}} \) are used (3). As with vaccinia virus infection, this yeast in vivo system shows that \( Z_{\text{ADHEL}} \) is active in binding Z-DNA. The experiments reported here suggest that residue Y48 undergoes a conformational change on binding Z-DNA in vitro. Thus, the change in the tyrosine side chain conformation may act as a switch to turn on in vivo activity.

We thank Heidemarie Lerch and Eberhard Krause for matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry and Heike Nikolkenko and Michael Bienert for providing a CD spectrometer. We further thank T. D. Goddard, G. Cornilesu and F. Delaglio, and P. Günter for providing the software SPARKY 3.105, TALOS, and DYANA 3.1/MOLMOL 2.1-2, respectively. This work was supported by grants from the National Institutes of Health (to A.R.).