C h a p t e r 4

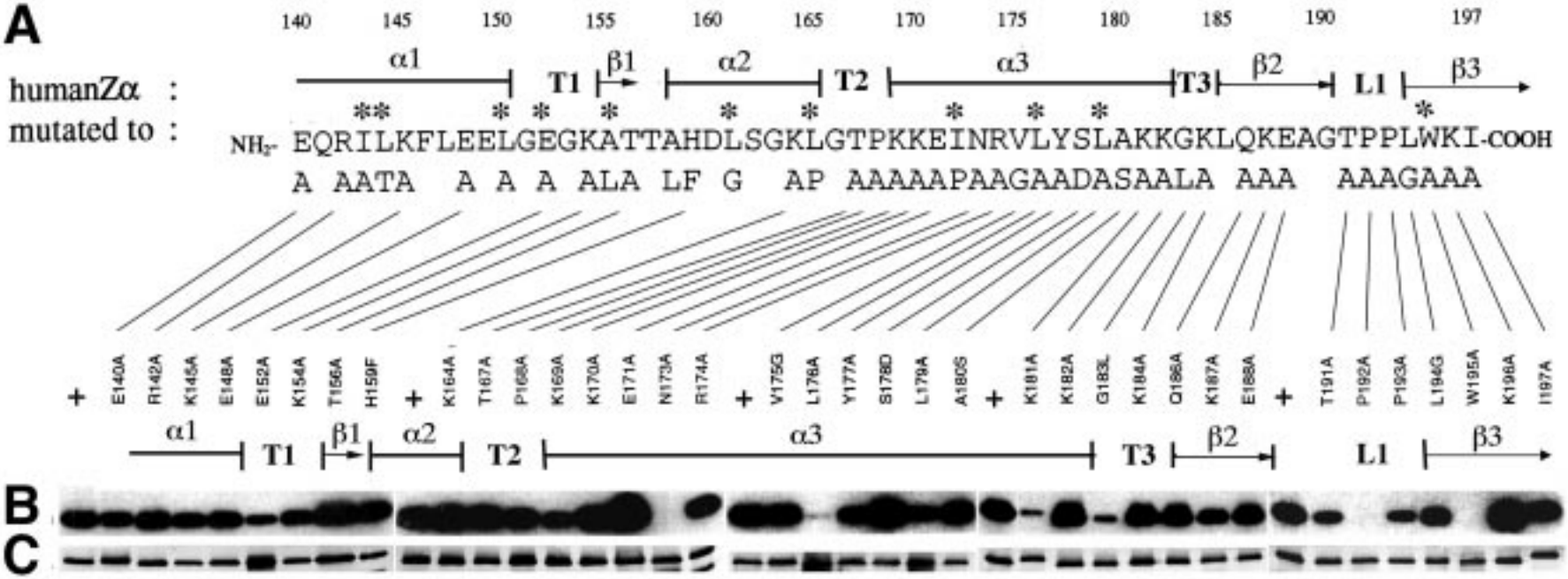
SCANNING MUTAGENESIS OF Zα

Mutagenesis and functional assays

At the time when the mutagenesis study was completed, the data on the secondary structure and topology of the Zα domain were available from solution-state NMR spectroscopy, as described in the next chapter. The NMR-derived topology (fig. 38) was set in a functional context by site-directed mutagenesis which is a technique allowing residue-specific mapping of protein interactions with DNA [172,173]. In most cases, alanine was substituted for the native residue. Alanine substitutions remove side chains beyond the ß carbon, thereby eliminating hydrophobic, salt-bridged and hydrogen-bridged side chain interactions, while maintaining the peptide backbone. In addition, other amino acids were used to replace alanines present in Zα, and also to test the effect of less conservative substitutions on Zα function. A total of 44 amino acids between residues 139 and 198 of the human Zα domain were systematically replaced using PCR-based site-directed mutagenesis (fig. 23A).

Mutant proteins were expressed as glutathione-S-transferase (GST) fusion proteins in *E.coli* and purified by affinity chromatography, allowing recovery of full length and proteolyzed mutant protein. Mutant proteins were resolved by SDS-PAGE, blotted to a nitrocellulose membrane and probed for Z-DNA binding with radioactively labeled Z-DNA (Southwestern assay) (fig. 23B). Only protein with a fully intact Zα domain bound Z-DNA (data not shown). Subsequent probing of the blot with anti-GST made it possible to measure the extent of C-terminal proteolytic degradation, an indirect measure of how a mutation affects protein stability *in vivo* (fig. 23C). Each mutant was also tested in a bandshift assay using the same DNA probe. The bandshift assay allows more accurate quantitation and can be used over a hundred fold greater dilution range than the southwestern. Dissociation constants (K_d) for the most deleterious mutants were measured with a BIAcore instrument under equilibrium conditions using an immobilized Z-DNA ligand as a substrate. Concordant results between assays were found except for some mutants discussed below.

fig. 23 **Site-directed mutagenesis of human Zα** (see next page). **(A)** Summary of mutants tested. Stars (*) indicate substantial (>80%) proteolytic degradation. **(B)** Southwestern assay. Autoradiogram of Z-DNA affinity assays, identifying those mutants that cause diminished Z-DNA binding. The panel is a collage of five different membranes. The lanes marked + are positive controls using wild-type Zα protein. **(C)** Western blot. Filters used for Southwestern assays were probed using antibodies to GST in order to quantitate protein loadings. Comparable band intensities show that equal amounts of full-length protein were assayed in C (Southwestern). Mutants sensitive to proteolysis, e.g. E152A, L176A and L179A, show a smearing downwards of the band, which results from overloading the SDS gel in order to ensure equal amounts of full length protein. In order to avoid extensive overloading of the gel, some mutants that exhibit severe proteolysis were excluded from the southwestern.


A
140 145 150 155 160 165 170 175 180 185 190 197
α1 T1 β1 α2 T2 α3 T3 β2 L1 β3
humanZα :
mutated to :
NH2- EQRILKFLEELGEGKATTAHDLSGKLGTPKKEINRVLYSLAKKGKLQKEAGTPPLWKI-COOH
A AATA A A A ALA LF G AP AAAAAPAAGAADASAALA AAA AAAGAAA
+ E140A R142A K145A E148A E152A K154A T156A H159F + K164A T167A P168A K169A K170A E171A N173A R174A + V175G L178A Y177A S178D L179A A180S + K181A K182A G183L K184A Q186A K187A E188A + T191A P192A P193A L194G W195A K196A I197A
α1 T1 β1 α2 T2 α3 T3 β2 L1 β3
B
C

Using the southwestern assay, we identified four groups of mutants in Zα that have effects on Z-DNA binding and protein stability; first, a cluster of mutants that lie in the C-terminal β-sheet; second, a collection of mutants in α3 that alter Z-DNA binding but not protein stability; third, a series of mutants that affect hydrophobic residues present in α1, α2 and α3 and disrupt protein stability; fourth, a number of mutants present in the turns between structural elements defined by NMR that reduce Z-DNA binding. A number of these mutants affect residues that are highly conserved between human, mouse, rat, bovine and *Xenopus* Zα sequences. Each set of mutants is discussed below.

The C-terminal β-sheet shows 2 distinct sides

The C-terminal β–sheet is composed of strands β2 and β3, connected by a four residue loop. The network of backbone NOEs between ß2 and β3 shows that the side chains of K187, W195 and I197 are next to each other, projecting to one side of the ß-sheet (side 1) The hydrophilic residues Q186, E188, and K196 project to the other side (side 2). Side chain NOEs show that side 1 of the β-sheet is in contact with α3.

Mutations to residues on side 1 of the β-sheet have profound effects. W195A is the most deleterious mutation of all Zα mutants made, regardless of the assay. Binding was diminished beyond detection and substantial proteolytic degradation was observed. These results confirm an important role for W195 in stabilizing Zα, and are consistent with the extensive long-range NOEs demonstrated for this residue by NMR (table 6, fig. 38). The more conservative W195Y mutant improved protein stability, but still caused a roughly 4-fold reduction in binding (bandshift assay, data not shown). In contrast, mutation of the other hydrophobic residue in β3, I197, to alanine had no effect on binding or protein stability. Changing the non-conserved residue K187 in β2 to alanine caused a mild but significant reduction in binding, raising the possibility that this flexible, positively charged side chain contacts DNA directly.

When residues on side 2 of the β-sheet were mutated, it was found that K196A enhanced Z-DNA binding while Q186A and E188A had wild-type protein stability and Z-DNA binding. Thus it appears that residues on this side of the β-sheet do not affect protein stability, but can indirectly influence DNA binding, for example, by modifying the dipole moment of the β-sheet.

The loop (L1) between β2 and β3 contains two highly conserved residues: P192 and P193. Mutant P192A had significantly diminished binding, while P193A showed only a slight reduction. As prolines often bend the protein backbone to form loops, the effects of these proline mutants may arise indirectly by disrupting β-sheet folding crucial for interaction with Z-DNA. Thus, P192 could be important in stabilizing the fold while P193 only makes minor contributions. However, a direct interaction between these residues and Z-DNA is also possible and cannot be excluded by this study. Mutation of the non-conserved L1 residue, T191, to alanine also diminished binding, suggesting that the T191 side chain may stabilize the loop structure, while L194G caused only a mild diminution in binding. Effects of mutating the other loop residue G190 were not tested. Thus of the four L1 residues mutated, all have some effect on DNA binding by Zα.

Helix α3 has characteristic properties of a recognition helix

Comparing results from NMR studies and alanine scanning mutagenesis shows that a striking cluster of deleterious mutations occurs in α3 (residues 169-182). An α-helical net diagram (fig. 24) segmenting α3 into three faces is useful in interpreting these mutants. Face 1 contains charged and polar residues. Mutations to these residues diminish Z-DNA binding but do not significantly affect protein stability. Face 2 is composed of hydrophobic residues. Changes to these amino acids are poorly tolerated, and cause increased protein degradation within *E. coli*. Face 3 is occupied by residues that do not reduce DNA-binding or affect protein stability. In some cases, mutation to these residues enhances Z-DNA binding. The three functionally distinct faces are discussed separately.

Residues K169, K170, N173, Y177 and K181 form face 1 (fig. 24). Residues K169, N173 and Y177 show a high degree of conservation among the human, rat, bovine and frog Zα, suggesting that they have an important role in Zα function. Residue 181 (K in humans) is poorly conserved. In the southwestern and bandshift assays, mutant N173A displays the most dramatic decrease in binding of all α3 mutations. Mutants K169A and Y177A show only a minor effect on binding in the southwestern assay, but show significant loss of activity in the bandshift assay, while mutant K181A lacks activity in both the southwestern and bandshift assays. CD spectra [data not shown] of these mutants are similar to wild-type, suggesting that secondary structure is also similar. When analyzed using BIAcore measurements N173A reduces binding 160-fold, K169A 37-fold, Y177A 26-fold and K181A 4-fold relative to the wild-type (table 2). The Kd for K181A is not reduced as much as would be predicted from bandshift and southwestern assays. A possible reason for this discrepancy is suggested by inspection of the BIAcore sensorgram shown in fig. 25. During the dissociation phase of the

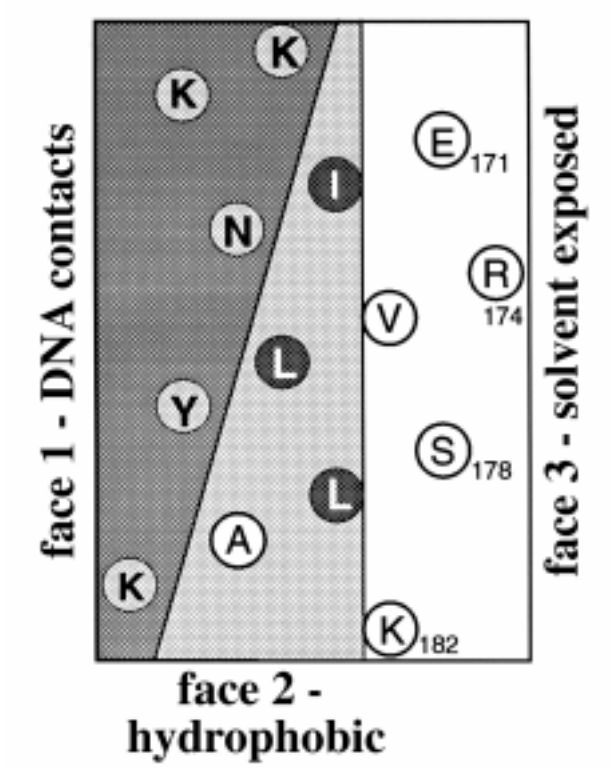


fig. 24 **α-helical net diagram of α3.** The putative DNA contacting face 1, the hydrophobic face 2 and the proposed solvent exposed face 3 are colored dark-gray, light-gray and white, respectively. Mutations to K169, N173, Y177 and K181 affect DNA binding without affecting protein stability. Changes to I172, L176 and L179 greatly diminished protein stability. Both E171A and S178D slightly increased Z-DNA binding, while R174A, V175A and K182A were identical to wild-type.

measurement, K181A shows a faster drop in signal than wild-type protein. The rapid breakdown of protein-DNA complexes during electrophoresis may account for loss of band-shifting activity of K181A and similarly, breakdown during washing may explain the diminished binding seen in the southwestern assay. The alignment of K169A, N173A, Y177A and K181A mutants in a row on one face of a helix is reminiscent of DNA contacts commonly found in the recognition helices of HTH DNA binding proteins [174,175]. Therefore, the face 1 residues are good candidates to form direct or water-mediated DNA contacts.

Additional mutants were made to examine the nature of interaction of K181 and Y177 with DNA. A K181R mutant binds indistinguishably from wild-type, suggesting that K181 recognizes Z-DNA in a flexible, non-restrained mode, e.g. through hydrogen bonds with the phosphate backbone. This result perhaps explains why K181 is replaced by a charged residue such as arginine or glutamate in other family members. A Y177K mutant was studied because the side chain of lysine is flexible enough to reach any putative hydrogen bonding partner that interacts with the hydroxyl of tyrosine. Mutant Y177K bound better than Y177A, but still worse than wild-type, showing that lysine can mimic some but not all the binding contributions mediated by the tyrosine side chain. In addition to hydrogen bonding via its hydroxyl group,

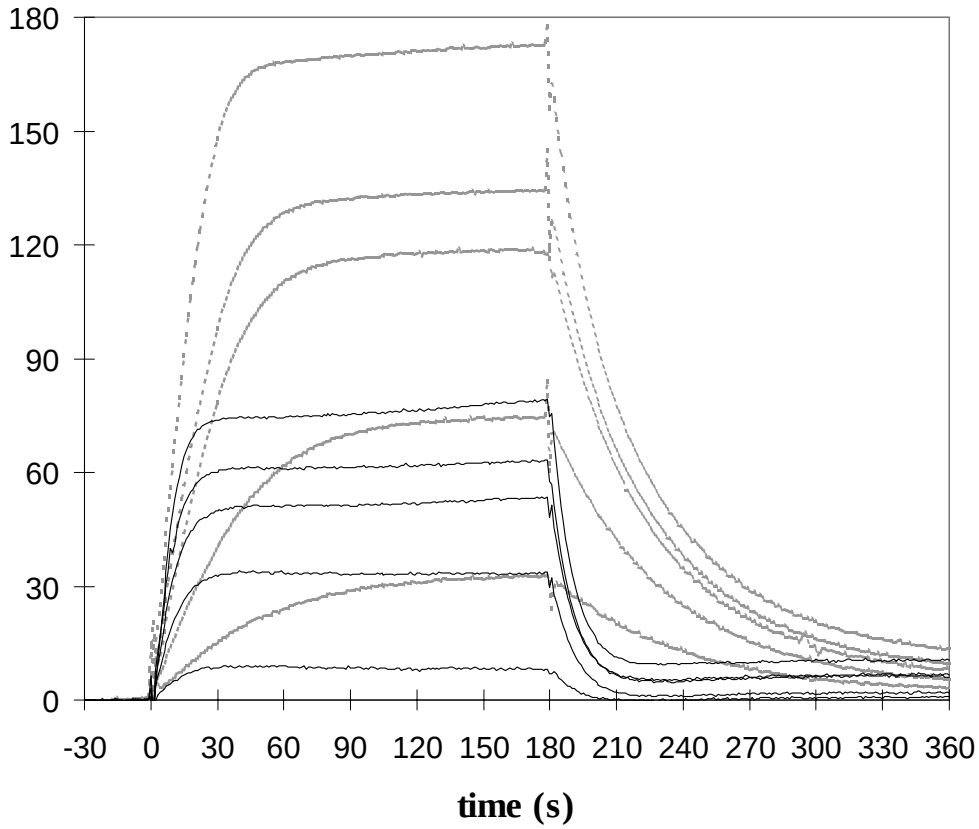


fig. 25 **BIAcore sensorgrams of wild-type and mutant K181A.** For wild-type Zα (dashed lines) the curves represent from top to bottom injections of 100, 75, 50, 25 and 10 nM of protein. For K181A, the curves (solid line) show injections of 100, 80, 60, 40 and 10 nM of protein. Response units (RU) are an arbitrary measurement that is proportional to the mass of protein bound to Z-DNA polymer immobilized on the chip surface. The more than 2-fold lower steady state response (plateau region) of mutant K181A reflects its lower affinity for Z-DNA. The rapid drop of response signal in the dissociation phase of K181A suggests a rapid breakdown of the DNA-protein complex.

Y177 is also likely to make essential contacts through its rigid aromatic ring system. The importance of hydrophobic contacts is suggested by the substitution of Y177 with another hydrophobic residue, isoleucine, in Zβ. Taken together, the mutational analysis revealed that residues K169, N173, Y177 and K181 are pivotal to binding, suggesting that face 1 of α3 interacts directly with Z-DNA and thus is a recognition helix.

The hydrophobic residues on face 2, I172, L176 and L179 are also highly conserved, with L176 being absolutely conserved between Zα and Zβ of all species. The result of mutating each of these residues was determined. A I172P mutation disrupted the protein fold so severely that no full length protein could be recovered from *E. coli*. A more conservative L176A mutation also caused extreme proteolytic degradation, allowing recovery of only 5% full length protein. Even after correcting for the reduced yield of full length protein, Z-DNA binding by this mutant was diminished 43-fold (table 2). Replacement of the third conserved residue, L179, with alanine results in slightly less protein degradation than L176A (~10% is full length protein) and Z-DNA binding assayed by southwestern and bandshift assays was somewhat better. These result show that mutation of hydrophobic residues on face 2 of α3 causes tremendous proteolytic degradation, suggesting that they form part of a hydrophobic protein core stabilizing Zα.

Face 3 of α3 carries three charged and one polar amino acid, none of which are conserved. Alanine mutants of the two positively charged residues R174 and K182 were indistinguishable from wild-type. The mutants E171A and S178D revealed a slightly enhanced Z-DNA affinity. Thus, these four mutants on face 3 are dispensable for Z-DNA binding and are probably solvent exposed.

table 2 **Binding affinities relative to wild-type**

Protein	relative Kd
wildtype	1
K169A	37
N173A	168
L176A	43
Y177A	26
K181A	4
P192A	13
W195A	>10000
engrailed	286

The Kd was determined by surface plasmon resonance (BIAcore) spectroscopy measuring the amount of protein bound to a Z-DNA coated chip surface after steady state was reached. Six mutants exhibit an affinity more than an order of magnitude lower than wild-type. Engrailed protein was used as a negative control for non-specific protein binding to Z-DNA. The binding by engrailed exceeded that of the most deleterious mutant W195A, for which a Kd could not be determined.

Helix α1 and α2 have essential hydrophobic faces

Mutation of α1 and α2 revealed conserved residues that form a hydrophobic face on each helix. In α1, the mutations I143A and L144T disrupt both binding and protein stability, while the L150A mutation increases proteolytic degradation without affecting binding. Together with L147, which was not investigated, these three residues make up a hydrophobic face on α1, suggesting that they are involved in contacts that correctly aligns this face within the protein. In contrast, mutation of α1 residues that lie on other faces of the helix, such as R142A, K145A and E148A, had no effect.

A similar face of hydrophobic residues is present in α2. Mutants A158L, L161G and L165P were found to be deleterious to both binding and protein stability. Replacement of A158 with a bulky leucine residue caused a milder phenotype than the severely destructive L161G and L165P mutations, suggesting that the environment of A158 at the joint of ß1 and α2 is less constrained than that of the two leucines. In summary, the hydrophobic faces of α1 and α2 probably contribute to a hydrophobic core essential for correct folding of Zα, similar to that found in other HTH proteins [174,175].

Essential turn residues

Another class of mutants maps to the turns deduced from the NMR secondary structure of Zα. T1 connects α1 and β1 and contains residues E152 and K154. Mutant E152A clearly has diminished binding and protein stability, suggesting that contacts formed by the glutamate side chain are essential for correct folding of Zα. NOEs that show contact between this residue and T156, which lies at the other end of β1 (table 6, fig. 38), are consistent with such a role for E152. The mutant K154A binds Z-DNA almost as tightly as wild-type. T2, between α2 and α3, consists of residues G166, T167 and P168, of which only the latter showed a minimal reduction in binding upon mutation to alanine, indicating that the main chain bend of P168 makes some contribution to the turn. Direct or water mediated contact of these side chains with DNA, as seen with other (α+β) HTH proteins [29], is unlikely. T3 residues G183 and K184 link the antiparallel β-sheet to α3. G183 is absolutely conserved in Zα and Zβ, while K184 is variable. The NMR data suggests that T3 must form a sharp turn to enable the ß-sheet to fold back against α3. Mutagenesis supports this suggestion. The introduction of a bulky leucine residue, which reduces the main chain flexibility, in place of G183 diminished Z-DNA binding substantially.

CD analysis of Y177A and K181A

To further understand the effect of Y177 and K181 on the interaction of Zα with Z-DNA, CD experiments were performed to test the effects of these mutants on the B-Z transition of polyd(CG). The CD spectra were collected after 10 minutes of incubation at 30° C for wild-type, K181A and Y177A at a molar ratio of 1 peptide to 1 basepair are shown in fig. 26. Also shown is the reference spectrum for the polymer in the B-DNA conformation obtained in buffer without any added protein. The spectra have not been corrected for absorption by the protein because this is absent above 250 nm and does not affect the interpretation of the DNA structural transition (data not shown). Below 250 nm, the signal from the peptide predominates. In fig. 26, the region between 240 and 320 nm is presented, showing the DNA transition from B-DNA to Z-DNA in the presence of protein [3]. The spectrum of Y177A shows a significantly reduced B- to Z-DNA transition compared to wild-type, while the K181A spectrum shows an enhancement. These results are somewhat paradoxical as it was expected that wild-type protein with the highest binding constant should be more effective in stabilizing Z-DNA than either of the mutants. However, CD is very sensitive to changes in the close environment of DNA, and these results may reflect slight variations in the Z-DNA conformation stabilized by each protein. Alternatively, K181A, which has only a 4-fold reduced affinity for Z-DNA (relative to Zα) and is used at high concentration in this assay (50 μM in fig. 26), may be more effective in promoting the B-Z transition of the polymer than wild-type.

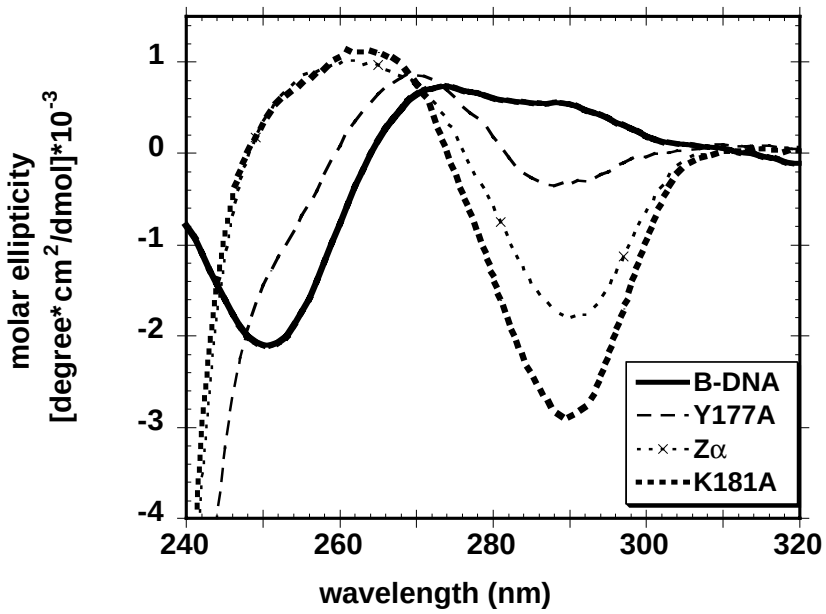


fig. 26 **Circular dichroism spectra showing the B-Z transition of polyd(CG) stabilized by wild-type and Y177A and K181A mutant Zα**. The CD titration of polyd(CG) with α3 face 1 mutants shows differences in their ability to promote formation of Z-DNA. The B-DNA spectrum of the polymer in the absence of protein is shown as a reference. The transition to Z-DNA in the presence of protein is shown by a loss of the negative peak at 250 nm and the appearance of a trough at 294 nm. Wild-type and Y177A and K181A mutant Zα were used at 1:1 = basepair:protein stochiometry (50 μM), and spectra were measured after a 10 minute equilibration period. The spectrum between 240 and 320 nm is presented to show changes in DNA conformation. Protein alone has a strong negative spectrum below 250 nm, but does not show a signal between 250 and 320 nm. Thus no correction for protein absorption has been made to the spectra shown.

Discussion

In this study, the topology of Zα derived from NMR experiments, was set in a functional context by scanning mutagenesis assigning functional motifs to various secondary structure elements. These data provide evidence that Zα belongs to the (α+β) helix-turn-helix family of DNA binding proteins, which show a similar arrangement of α-helices and β-sheets to that found in histone H5 (GH5) [176] and hepatocyte nuclear factor (HNF-3γ) [36]. A primary sequence alignment of Zα with HNF-3γ and histone H5 (fig. 27) shows that helical regions have a similar distribution of hydrophobic residues. The hydrophobic residues in Zα that were identified by mutagenesis to be essential for protein stability (fig. 23) align well with residues in histone H5 and HNF-3γ that were shown by X-ray crystallography to lie within the hydrophobic protein core (fig. 27).

Furthermore, the mutagenesis data suggests that α3 of Zα acts as a recognition helix for Z-DNA, playing a role analogous to that of α3 in HNF-3γ and GH5 in the recognition of B-

DNA. Zα, HNF-3γ and GH5 also appear to have C-terminal double-stranded β-sheets that are similar in topology. In Zα and HNF-3γ, a conserved tryptophan present in β3 makes contact with α3. By analogy with HNF-3γ and consistent with the mutagenesis data presented here, the tryptophan of Zα may be hydrogen bonded with DNA phosphates through its imino group. The extensive contacts made by W195 shown in fig. 38 indicate that this residue is also essential for stabilizing the Zα fold. All of the above results suggest that the global protein fold of Zα resembles that of (α+β)HTH family members HNF-3γ and histone H5.

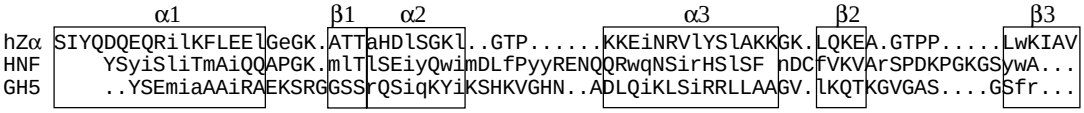


fig. 27 **Primary sequence alignment of Zα with HNF-3γ (HNF) and histone H5 (GH5)** using the preliminary secondary structure of Zα (fig. 35) and the crystal structures of HNF-3γ [36] and histone H5 [176]. Zα residues in lower case letters greatly affect protein stability (mutation of these residues results in less than 20% recovery of full-length protein from *E. coli*). HNF-3γ and histone H5 residues shown in lower case letters are buried amino acids contributing to the hydrophobic protein core.

Each member of the (α+β)HTH family has a unique orientation of the recognition helix to DNA, arising from differences in primary sequence, helix length, the number of residues in loops and the composition of β sheets. As a consequence, this family of proteins can bind a wide variety of sequences and conformations, including bent DNA [37,177,178]. The primary sequence of family members is quite divergent. For example, histone H5 and HNF-3γ lack significant sequence homology, having identity in only 9 of 72 residues [36]. However, their Cα backbone shows a minor root mean standard deviation of 1.3 Å. Similarly, Zα is not closely related to either histone H5 or HNF-3/*fork head* proteins at the primary sequence level (fig. 27), even though they appear to share a common protein fold. Zα has many other significant differences from these two proteins. These differences are most apparent when it is compared to HNF-3γ, for which the structure bound to DNA is available [36]. For example, the loop in HNF-3γ between α2 and the recognition helix α3 is eight residues while the turn T2 of Zα is only three. This loop in HNF-3γ is important in determining sequence specificity and for the relative orientation of α3 relative to α2 [29], but there is no evidence that the side chains of T2 residues in Zα are involved in DNA recognition. In addition, binding by HNF-3γ to DNA involves a single residue in the eight residue loop between β2 and β3 (called W1, equivalent to L1 in Zα) and a number of residues in the carboxy-terminal loop, W2, which extends beyond the C-terminal β-sheet. Zα does not appear to require a region equivalent to W2 for tight binding to Z-DNA while mutation of three of the four residues in L1 (G190 was not tested) affect DNA binding. Furthermore, in Zα, both the amino and carboxy termini of the putative recognition helix, α3, have lysine residues that affect binding to Z-DNA, whereas in HNF-3γ there is no equivalent. The above differences between Zα and HNF-3γ are consistent with their distinct modes of interaction with DNA.

It is possible that (α+β)HTH motifs recognize Z-DNA in addition to B-DNA. A wide range of biological processes are regulated by the (α+β) family of HTH proteins, such as embryogenesis, tumorigenesis, longevity, insulin-related metabolic pathways and maintenance of differentiated cell states [29,179]. The relationship of Zα to this family of proteins raises the

intriguing question of whether other members of this family may also bind to the left-handed DNA conformation.

Summary

RNA editing alters pre-mRNA through site-selective adenosine deamination, which results in codon changes that lead to the production of novel proteins. An enzyme that catalyzes this reaction, double-stranded RNA adenosine deaminase (ADAR1), contains two N-terminal Z-DNA binding motifs, Zα and Zβ, of as yet unknown function. Here, site-directed mutagenesis was used to identify residues in α3, in β3 and the loop connecting β2 to β3, that diminish Z-DNA binding when mutated. Also identified were 11 hydrophobic residues that are essential for protein stability. These results were set in a structural context using the data on the topology of Zα derived from multidimensional NMR spectroscopy. Comparison with known structures reveals some similarity between Zα and (α+β) helix-turn-helix proteins, such as histone H5 and the family of HNF-3 (α+β) helix-turn-helix transcription factors. Taken together, the structural and functional data suggest that recognition of Z-DNA by Zα involves residues in both helix α3 and the C-terminal β-sheet.