Leucine Zipper

Dmitry Krylov, *National Cancer Institute, Bethesda, Maryland, USA*
Charles R Vinson, *National Cancer Institute, Bethesda, Maryland, USA*

The leucine zipper is a protein–protein interaction domain consisting of amphipathic α helices that dimerize in parallel, either as homodimers or heterodimers, to form a coiled-coil.

Secondary article

Article Contents

- Introduction
- Sequence Rules for Formation
- Structural Examples
- Stabilizing Forces: Role of Hydrophobic Effect, Side-chain Packing Interactions, Salt Bridges and Length
- Specificity of Interactions: Homo- versus Heterodimers
- Functional Example: Transcription, Interaction of Leucine Zipper Proteins with DNA, Interactions with Regulatory Proteins

Introduction

The leucine zipper is the dimerization domain of the B-ZIP (basic-region leucine zipper) class of eukaryotic transcription factors (Vinson *et al*., 1989). The name arose because leucines occur every seven amino acids in this dimerization domain. These leucines are critical for the dimerization and DNA binding of B-ZIP proteins. The leucine zipper is a left-handed parallel dimeric coiled-coil, a structure proposed independently by Pauling and Corey, and by Crick in 1953.

Sequence Rules for Formation

Figure 1 represents the X-ray structure of a B-ZIP dimer (GCN4) bound to DNA (Ellenberger *et al*., 1992) Each monomer is a long bipartite α helix. The N-terminal basic region contains the amino acids lysine and arginine, which interact with the major groove of DNA in a sequence-specific manner. The C-terminus is an amphipathic α helix that dimerizes to form the leucine zipper. An amphipathic α helix has two surfaces along its length. One side is hydrophobic (water-hating) and dimerizes with itself or another amphipathic α helix to remove the nonpolar amino acids from the water. The other side of the amphipathic α helix is hydrophilic (water-loving) and interacts with water. The burying of hydrophobic amino acids is the major physical driving force that produces a dimeric structure. The solvent-exposed surface of the dimer is hydrophilic. The protein sequence of the leucine zipper is composed of heptad (seven) repeats of amino acids, which can be added together to generate structures of varying length. In order to generate a repeating helical dimerization interface from the two sides of the amphipathic α helices in the dimer, the right-handed α helix over-twists slightly, effectively reducing the number of residues from 3.6 per turn in an amphipathic α helix to 3.5 amino acids per turn in the leucine zipper coiled-coil. The twisting around each other produces a repeating structure every two α-helical turns or seven amino acids (a heptad repeat). The hydrophobic surface of each amphipathic α helix in the dimer is composed of hydrophobic amino acids with a 3–4 pattern.

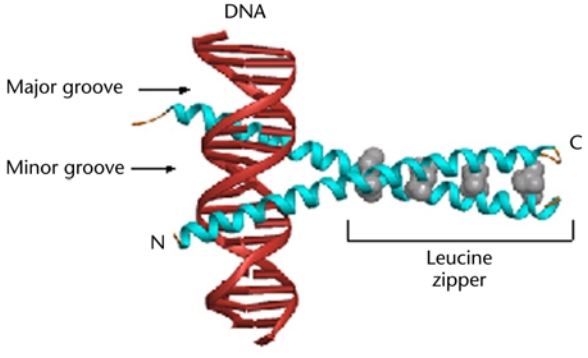

Figure 1 X-ray structure of the B-ZIP dimer GCN4 bound to DNA. The DNA is in red, the α helices are in blue. The d or leucine position amino acids are shown in grey. The N-terminal and C-terminal parts of the protein are labelled N and C.

Figure 2 shows a schematic of the side and end views of the coiled-coil structure using the standard nomenclature for the seven unique amino acid positions (a, b, c, d, e, f and g) in a heptad. Amino acids on the opposite helix of the dimer are designated a′, b′, c′, d′, e′, f′ and g′. The a and d residues are hydrophobic and pack in a regular 'knobs and holes' pattern along the dimerization interface to create the hydrophobic core. This contributes most of the energy that stabilizes dimerization. The regular spacing of hydrophobic amino acids in a 3–4 stutter pattern (a–d(3);d–a(4)) is critical for the formation of this structure. The 'knobs and holes' packing, which is the hallmark of a coiled-coil, means that the side-chain in the a or d position (knob) packs into the space surrounded by four side-chains (hole) on the opposite helix. The a side-chain is packed into the space surrounded by two d′, an a′, and a g′ side-chain. The d side-chain is packed into the space surrounded by two a′, a d′, and an e′ side-chain. The e and g positions, which flank the dimerization interface, often contain charged amino acids that interact interhelically to repel or form attractive g↔e′ salt bridges. These charge–charge interactions are important in the regulation of dimerization specificity.

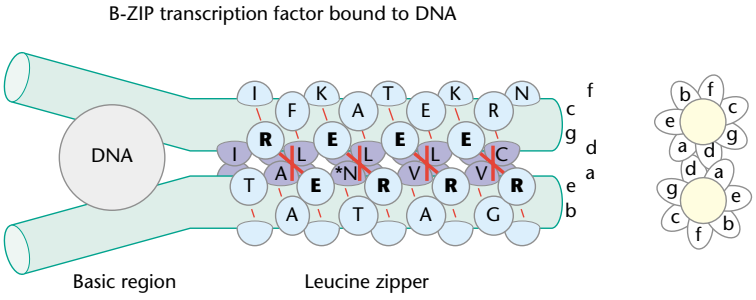

Figure 2 A schematic of the B-ZIP PAR family member VBP viewed from the side with the amino acids from the VBP leucine zipper shown inside the circles which represent amino acid positions along the two α helices. Amino acids in the e and g position are shown in bold face and the *i*, *i*′ + 5 (g↔e′) interactions are connected by arrows pointing from acidic to basic. The heptad letter designations (a, b, c, d, e, f, g) are shown. The supercoiling of the two helices is not depicted. To the left of the leucine zipper is the basic region of B-ZIP proteins with the DNA shown. To the right is an end view of a leucine zipper dimer looking from the N-terminus. The letters on the inside of each ellipse represents the standard nomenclature for the seven amino acids found in unique positions in a coiled-coil. The ellipses depict the orientations of the amino acid side-chains relative to the α helix. Amino acids in the a and d positions create a hydrophobic core between the interacting helices. The interaction seen between amino acids in the g and subsequent e′ position seen in X-ray structures is noted as g↔e′ pairs. Note that because of the 2-fold symmetry of the dimers, each heptad contains two g↔e′ pairs.

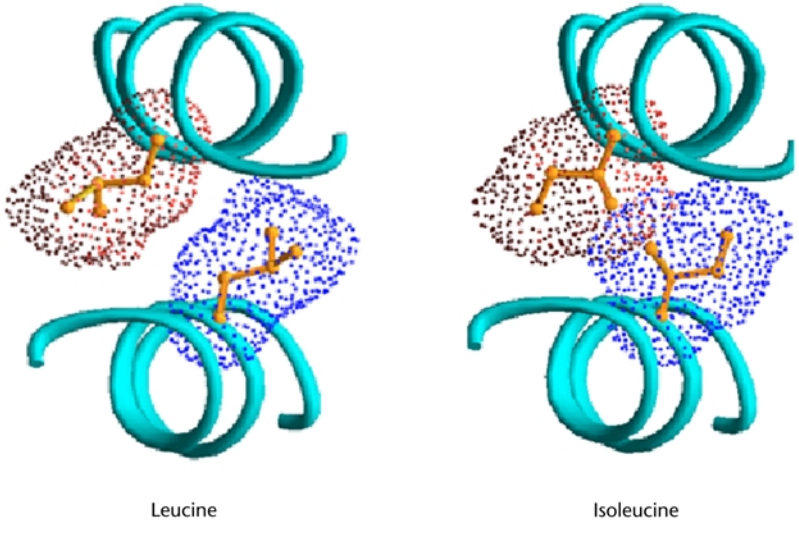

Figure 3 An end view, looking from the N-terminus, of the leucine zipper interface with either leucine or isoleucine in the d and d′ positions. The clockwise blue-green spirals represent α helices. The space-filling dots represent the volume of the side-chains. Note that the leucines pack nicely together while the isoleucines overlap, which is not possible physically.

The leucine zipper is a dimeric parallel coiled-coil, but amphipathic helices can also oligomerize to form parallel coiled-coils that are trimers, tetramers or pentamers. The majority of B-ZIP leucine zippers contain valines in the a position and leucines in the d position. Kim and colleagues changed both of these amino acids to isoleucine, which caused the leucine zippers to form trimers rather than dimers (Harbury *et al.*, 1993). This result can be appreciated from a structural viewpoint by examining the preferred side-chain rotamers of these two amino acids when placed in the d and d′ position of the GCN4 leucine zipper X-ray structure. Leucine is not a β-branched side-chain and can assume a variety of rotamers, one of which is able to pack well into the hydrophobic interface of the dimer. Isoleucine, in contrast, is β branched and only has one favoured side-chain rotamer for the Cβ carbon to avoid steric clash with the α helix. This allowed rotamer does not pack into the hydrophobic core of the dimer (**Figure 3**), forcing the hydrophobic core to open up and form a trimer. These studies, which systematically varied the a and d positions, have enhanced our understanding of amphipathic helix interactions but cannot explain the full range of oligomers observed.

Structural Examples

Fifty years ago, data from the X-ray diffraction of the dimeric coiled-coil fibrous proteins, which include keratins, myosins and fibronectin, were used to predict the coiled-coil structure. A two heptad 'trigger' sequence is critical to initiate the folding of long coiled-coils (Kammerer *et al*., 1998; Steinmetz *et al*., 1998). The identification of the leucine zipper domain in B-ZIP transcription factors reignited interest in this structure, as these structures were small and experimentally tractable. The first X-ray structure of the leucine zipper was determined in 1991 from the yeast GCN4 protein (O'Shea *et al*., 1991). The stunning conclusion was that the structure was exactly as predicted by Crick. The hydrophobic amino acids in the a and d position pack in the characteristic 'knobs and holes' pattern with a 3–4 repeating stutter. The long charged amino acids in the e and g positions lie across the hydrophobic interface and appear to interact favourably. This interhelical interaction is involved in regulating the dimerization specificity of amphipathic α helices. Crystal structures of the B-ZIP protein GCN4 bound to DNA show very similar structures for the leucine zipper region. The Fos and Jun heterodimer also contains the canonical 'holes and knobs' packing. An unexpected property of the Fos/Jun heterodimer bound to DNA was that Fos zipper was relatively straight while the Jun zipper wrapped around the Fos zipper. The amino acids regulating this asymmetric interaction between Fos and Jun remain obscure.

An impressive example of a coiled-coil structure undergoing a physiological structural transition comes from the work on the trimeric coiled-coil in the influenza haemagglutinin receptor (Carr and Kim, 1993). At physiological pH (pH 7), the receptor has a short trimeric coiled-coil with the N-terminal sequence forming α helices that lie on the outside of the trimer. At pH 4, which occurs in the lysosome after viral endocytosis, the trimeric coiled-coil extends N-terminally. The N-terminal helices lying on the outside of the coiled-coil at pH 7 on the surface of the cell now form an extension of the coiled-coil. An examination of the amino acids in this sequence indicates that the N-terminal section contains many glutamic acids. Glutamic acids are incompatible with a coiled-coil structure at pH 7 because of repulsive interactions between themselves, but at pH 4 the repulsion is neutralized so that the sequence can form an α helix that extends the trimeric coiled-coil.

Stabilizing Forces: Role of Hydrophobic Effect, Side-chain Packing Interactions, Salt Bridges and Length

The contribution of a particular amino acid to the stability of a leucine zipper has been determined by producing mutant proteins with only a single amino acid changed. The energetic consequence of each change has been determined by measuring the thermal stability of the new protein using circular dichroism (CD) spectroscopy. This technique measures the ellipticity of the sample, which is a reflection of the α-helical content. The leucine zipper motif can be denatured reversibly in a simple two-state manner using either heat or the denaturants urea or guanidine. This allows the application of rigorous thermodynamic calculations to the experimental data. The contribution of aliphatic amino acids in the fourth d position of VBP (**Figure 2**) to stability are shown in **Table 1**. CD experiments indicate that the hydrophobic amino acids in the a and d positions account for most of the stabilizing energy of this structure, e.g. leucine in the d position contributes

Table 1 Estimation of the van der Waals contacts for amino acids in the d position (kcal mol^{-1})

Amino acid	$\Delta\Delta G_A$	$\Delta\Delta G_{helix}$	$\Delta\Delta G_A - \Delta\Delta G_{helix}$	$\Delta\Delta G_{transfer}$	$\Delta\Delta G_{packing}$
Leu	–4.6	+0.15	–4.75	–2.32	–2.85
Met	–2.0	+0.27	–2.27	–1.68	–1.01
Ile	–1.65	+0.54	–2.19	–2.46	–0.25
Val	–1.1	+0.63	–1.73	–1.66	–0.49
Cys	–0.95	+0.54	–1.49	–2.10	+0.19
Ala	0.0	0.0	0.0	–0.42	0.0
Ser	+0.25	+0.42	–0.17	+0.05	–0.64

$\Delta\Delta G_A$ is the free energy of unfolding for a single amino acid in the monomer relative to alanine. $\Delta\Delta G_{helix}$ is the helical propensity of amino acids relative to alanine. $\Delta\Delta G_{transfer}$ is the change in the solvent transfer free energy relative to alanine. The difference between $\Delta\Delta G_A$ and the sum of $\Delta\Delta G_{transfer}$ and $\Delta\Delta G_{helix}$ is defined as the $\Delta\Delta G_{packing}$.

From Moitra *et al*. (1997).

9.2 kcal mol^{-1} per dimer relative to alanine (Moitra *et al.*, 1997). Of this, approximately half is from the transfer free energy, revealing the hydrophobic nature of leucine. The other half is from the enthalpic van der Waals interactions or packing of the leucine with other amino acids in the hydrophobic spine of the leucine zipper.

Experiments examining the e and g positions indicate that hydrophobic amino acids are more stabilizing than the charged amino acids that are typically in these positions. Hydrophobic amino acids in the e and g positions tend to cause amphipathic α helices to form higher order oligomers. The long charged amino acids (glutamic acid, lysine and arginine) frequently found in the e and g positions prevent higher order oligomers from forming and are critical for regulating the specificity of dimerization.

Specificity of Interactions: Homo- versus Heterodimers

Over 60 B-ZIP proteins have been identified, but an understanding of their dimerization preferences is in its infancy. Experiments with natural B-ZIP proteins indicate that some prefer to homodimerize and others prefer to heterodimerize. Using the B-ZIP heterodimerizing system of Fos and Jun, Kim and workers found that amino acids in the e and g positions were critical for dimerization specificity and amino acids in the a and d positions had little effect. (O'Shea *et al.*, 1992). More detailed work by Vinson and Hodges identified charged amino acids interacting between the g position of one helix and the following e′ position of the opposite helix (Krylov *et al.*, 1994; Zhou *et al.*, 1994). The general conclusion is that oppositely charged amino acids are attractive and like charged amino acids are repulsive.

The specificity of any protein–protein interaction can be characterized by the energetic interaction between individual amino acids. One method used to determine the energetics of interaction between any two amino acids in a protein structure is a double mutant thermodynamic analysis. **Figure 4** presents this analysis for the i, $i' + 5$ (g↔e′) interaction between glutamic acid in the g position and arginine in the following e′ position. This analysis consists of examining four proteins. The first protein contains alanines in both the g and the following e′ positions (A↔A). Alanine, consisting of a single methylene, is essentially a truncated amino acid. The second protein contains glutamic acid (E↔A) and the third protein contains arginine (A↔R). The fourth protein contains both glutamic acid and arginine (E↔R). The stability of the second, third and fourth proteins is compared with that of the first protein containing two alanines. If the two amino acids do not interact energetically, the stability of the fourth protein containing both amino acids will simply be the sum of the energy

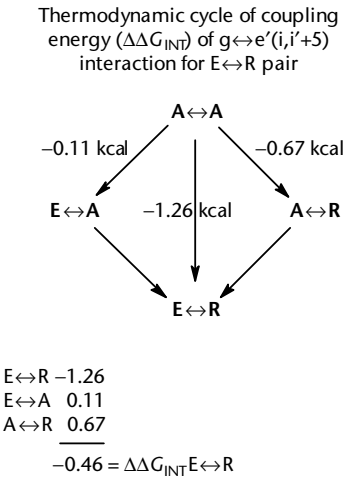

Figure 4 Double mutant alanine thermodynamic cycle used to determine coupling energy ($\Delta\Delta G_{int}$) for the interaction of glutamic acid (E) in the g position with arginine (R) in the following e′ position. The $\Delta\Delta G$ values presented are in terms of an individual g↔e′ interaction. The E↔R pair is 1.26 kcal mol^{-1} more stable than the A↔A pair. The contribution of the individual amino acids to the stability of the leucine zipper was determined by studying proteins containing only glutamic (E↔A) or arginine (A↔R) of the pair. The E↔A pair is 0.11 kcal mol^{-1} more stable than the A↔A pair. The A↔R pair is 0.67 kcal mol^{-1} more stable than A↔A. The sum of the individual contributions of E and R to the dimer stability is -0.78 kcal mol^{-1}. The extra -0.46 kcal mol^{-1} of stability ($-1.26 - (-0.78)$) from the E↔R pair is the coupling energy ($\Delta\Delta G_{int}$), indicative of the interaction of E with R across the surface of the leucine zipper.

contribution of each amino acid. However, if the two amino acids interact positively, the protein containing both will be more stable than the sum of the contribution of the individual amino acids. Conversely, if they interact repulsively, it will be less stable than expected.

The four most common amino acids in the g and e positions of the leucine zippers of B-ZIP proteins contain long side-chains to reach across the hydrophobic interface. They include the two basic amino acids, lysine and arginine, the acidic amino acid glutamic acid, and the polar amino acid. A double mutant thermodynamic analysis of these four amino acids was used to determine the energetic contribution (**Table 2**: ΔG) and the coupling energy (**Table 3**: $\Delta\Delta G$) between these amino acids in the g and e positions (Krylov *et al.*, 1998). **Figure 2** shows the two g↔e′ pairs of VBP that were mutated. The results show that the basic amino acids, particularly arginine, interact attractively with glutamic acid with an energetic contribution of 1.3 kcal mol^{-1} and a coupling energy of -0.5 kcal mol^{-1}. Only the glutamic acid interaction (E↔E) is less stable ($+0.4$ kcal mol^{-1}) than the alanine–alanine interaction (A↔A). The coupling energy between glutamic acid (E↔E) is repulsive ($+0.8$ kcal mol^{-1}). The calculation of coupling energy is illustrated in **Figure 4**. These results alone do not explain the dimerization preferences of B-ZIP

Table 2 Thermodynamic differences for *i*, *i*'+5, (g↔e') interactions relative to A↔A (ΔG_{AA}) (kcal mol^{-1} per salt bridge)

↓ g\e'→	A	E	Q	R	K
A	0.0	–0.25	–0.75	–0.67	–0.54
E	–0.11	+0.38	–0.73	–1.26	–0.97
Q	–0.39	–0.46	–1.17	–0.79	–0.83
R	–0.21	–1.55	–0.58	–0.10	–0.10
K	–0.24	–1.42	–0.71	–0.32	–0.34
D	>+1.57	–	–	+0.38	+0.62

Circular dichroism (CD) thermal denaturation was monitored at 222 nm in 150 mmol L^{-1} KCl and 12.5 mmol L^{-1} phosphate pH 7.4.
From Krylov *et al.* (1998).

Table 3 Coupling energy of interaction ($\Delta\Delta G$int) for *i*, *i*′ + 5 (g↔e') pairs (kcal mol^{-1} per salt bridge)

↓g\e′→	E	Q	R	K
E	+ 0.8	+ 0.16	− 0.5	− 0.3
Q	+ 0.20	− 0.03	− 0.35	+ 0.26
R	− 1.07	+ 0.38	+ 0.16	+ 0.81
K	− 0.91	+ 0.28	− 0.03	+ 0.60

Values were calculated from **Table 2**. See **Figure 4**.

proteins, which suggests the presence of additional interactions regulating dimerization specificity.

This combination of attractive and repulsive interactions between amphipathic helices results in a set of dimerization possibilities. These features can be used to design helices with specific dimerization properties. A protein can be designed with attractive electrostatic interactions that will only homodimerize and not interact with other B-ZIP monomers. Alternatively, a stable protein can be built with repulsive homodimer interactions that will promiscuously interact with many other B-ZIP monomers. The ability to change dimerization partners allows for a complex set of possible heterodimers to form. For example, one can have a B-ZIP protein that homodimerizes, binds DNA and regulates gene expression but this can be changed by expressing a new B-ZIP protein that preferentially heterodimerizes with the initial protein. Now only heterodimers will form, bind to a new sequence of DNA and regulate expression of new genes. This hypothetical example helps in understanding the complex set of genes that can be regulated by simply having different concentrations of B-ZIP monomers.

Functional Example: Transcription, Interaction of Leucine Zipper Proteins with DNA, Interactions with Regulatory Proteins

The B-ZIP monomer consists of a long α helix. The C-terminal half is the leucine zipper and the N-terminal half binds to the major groove of DNA in a sequence-specific fashion (**Figure 1**). A conserved arginine and a conserved asparagine in the basic region are exactly positioned relative to the leucine zipper, indicating the importance of the juxtaposition of these two structural elements for effective DNA binding (Vinson *et al.*, 1989). In the absence of DNA, the DNA-binding region is not structured, but upon DNA binding, it becomes α helical, lying in the major groove of DNA. Each helical extension of the leucine zipper can bind up to 5 base pairs in a sequence-specific manner. Thus, the dimer can bind up to 10 base pairs without crossing the DNA backbone. This form of DNA binding is dramatically different from most DNA-binding motifs. In the prototypical helix-turn-helix DNA-recognition motif, one helix lies in the major groove of DNA and a second helix lies across it, interacting with the DNA backbone and essentially pinning the DNA recognition

helix in place. In B-ZIP proteins nothing holds the DNA recognition helix in place except that it is oriented by being an extension of an α helix protruding from a leucine zipper.

Typically, a homodimer binds an abutted palindromic DNA sequence. Heterodimers can bind any combination of half sites. The rules regulating DNA-binding specificity are poorly understood for this class of proteins. For the B-ZIP dimer to bind DNA, the leucine zipper has to interact in parallel and in heptad register to place both basic regions in the major groove. One structural feature of the leucine zipper that accomplishes this heptad register is a nearly invariant asparagine (the Fos family has a lysine in this position) in the a position of the leucine zipper of B-ZIP proteins. These asparagines can form hydrogen bonds with each other in the hydrophobic core. When leucine zippers are out of phase with each other, the asparagines become juxtaposed with a hydrophobic amino acid, making the interaction energetically unstable. This is a classic example in which possible protein conformations are limited by placing a hydrophilic amino acid (asparagine) in a hydrophobic core.

To assist in unravelling the biological function of a B-ZIP protein, it is useful to use proteins that inhibit the DNA binding and function of individual B-ZIP genes. These types of proteins are called dominant-negatives. The term 'dominant' refers to their genetic dominance. The term 'negative' describes the inhibition of the function of cellular proteins. In its simplest form, a dominant-negative could be a truncated B-ZIP protein, which would heterodimerize with endogenous B-ZIP proteins, producing an inactive heterodimer. However, the binding of DNA stabilizes the B-ZIP structure and this complicates the design of biologically active dominant-negative proteins. This complication has been overcome by designing dominant-negatives based on the concept that an amino acid sequence could mimic the properties of DNA. Knowing that the B-ZIP basic region could form an α helix when bound to DNA, a protein sequence was designed with the hope that it could mimic DNA. In these dominant-negatives, termed A-ZIPs (A for acidic extension and the ZIP for leucine zipper), the designed protein sequence replaces the DNA-binding region with a designed acidic amphipathic helix. This protein sequence contains a hydrophobic core in the a and d positions and acidic amino acids in the g and e positions. The acidic extension forms a coiled-coil with the basic region of the wild-type B-ZIP protein, essentially extending the leucine zipper into the basic region. This stabilizes the heterodimer complex and prevents the B-ZIP complex from binding DNA. The stability achieved by extending the leucine zipper is essential to form a dominant-negative that can compete with DNA binding. This type of dominant-negative reveals both the activating and the repressing properties of a particular family of B-ZIP transcription factors.

An examination of amino acids in the leucine zipper region of B-ZIP proteins indicates that all seven positions

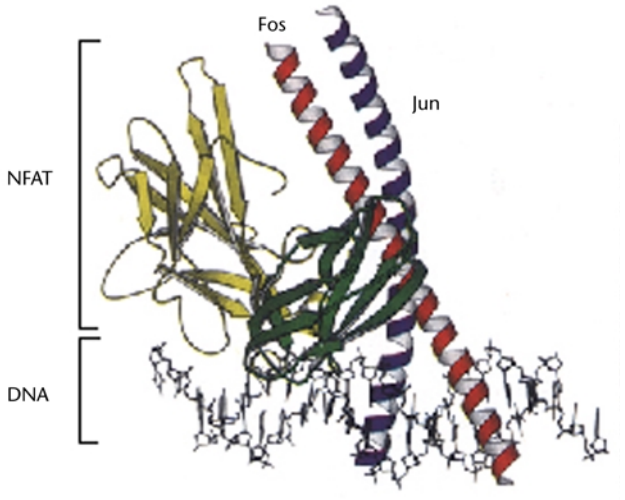

Figure 5 X-ray structure showing the Fos/Jun heterodimer with the basic region bound to DNA and the leucine zipper region interacting with another DNA-binding protein, nuclear factor of T cells (NFAT). Interactions with NFAT occur through conserved amino acids in the b, c and f positions of the Fos/Jun leucine zipper.

are conserved. So far, we know that the a and d positions are important for dimer stability and e and g positions are important for dimerization specificity. Little is known about the importance of the conserved amino acids in the b, c and f positions that are on the outside surface of the leucine zipper. Their conservation suggests that additional proteins may interact with coiled-coil dimers at the b, c and f positions. The best molecular example of a B-ZIP protein interacting with another protein via the b, c and f positions is the co-crystal of the DNA-binding domains of Fos and Jun bound to specific DNA, interacting with another DNA-binding protein, nuclear factor of T cells (NFAT) (**Figure 5**). This structure describes interactions between the b, c and f positions of the coiled-coil, and the NFAT protein. In the future, we should see additional examples of B-ZIP dimers bound to DNA, and B-ZIP dimers interacting with other DNA-binding proteins on DNA.

References

Carr CM and Kim PS (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell* **73**(4): 823–832.

Crick F (1953) The packing of Ã-helices: simple coiled-coils. *Acta Crystallographica* **6**: 689–697.

Ellenberger TE, Brandl CJ, Struhl K and Harrison SC (1992) The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein–DNA complex. *Cell* **71**(7): 1223–1237.

Harbury PB, Zhang T, Kim PS and Alber T (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**: 1401–1407.

Kammerer RA, Schulthess T, Landwehr R *et al.* (1998) An autonomous folding unit mediates the assembly of two-stranded coiled coils. *Proceedings of the National Academy of Sciences of the USA* **95**: 13419–13424.

Krylov D, Mikhailenko I and Vinson C (1994) A thermodynamic scale for leucine zipper stability and dimerization specificity: E and G interhelical interactions. *EMBO Journal* **13**: 1849–1861.

Krylov D, Barchi J and Vinson C (1998) Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *Journal of Molecular Biology* **279**: 959–972.

Moitra J, Szilák L, Krylov D and Vinson C (1997) Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **36**: 12567–12573.

O'Shea EK, Klemm JD, Kim PS and Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**: 539–544.

O'Shea EK, Rutkowski R and Kim PS (1992) Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell* **68**: 699–708.

Steinmetz MO, Stock A, Schulthess T *et al*. (1998) A distinct 14 residue site triggers coiled-coil formation in cortexillin I. *EMBO Journal* **17**: 1883–1891.

Vinson CR, Sigler PB and McKnight SL (1989) Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* **246**: 911–916.

Zhou N, Kay C and Hodges R (1994) The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Engineering* **7**(11): 1365–1372.