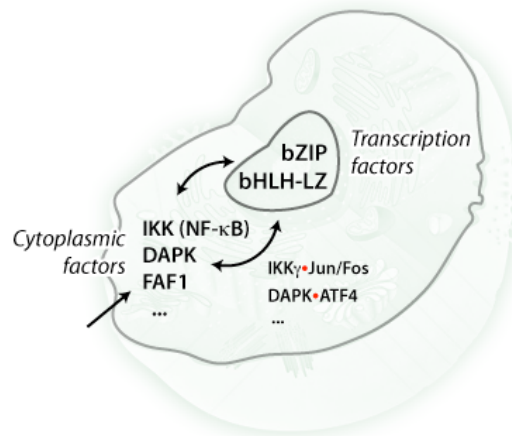


# Rethinking Leucine Zipper – a universal signal transduction motif

Yaroslav Nikolaev and Konstantin Pervushin

*Biozentrum of University Basel, Klingelbergstrasse 70, CH-4056, Basel, Switzerland*

mailto: [y.nikolaev@unibas.ch](mailto:y.nikolaev@unibas.ch)



This work is supported by the Swiss National Science Foundation Grant 3100A0-118381.

# Table of contents

<b>= 1 = Introduction</b>	<b>4</b>
<b>= 2 = Structure</b>	<b>5</b>
= 2.1 = Primary – heptad repeat	5
= 2.2 = Secondary and tertiary – stability and stoichiometry	6
= 2.3 = Quaternary – specificity	8
<b>= 3 = Stability and specificity</b>	<b>8</b>
= 3.1 = D-D' interactions (stability)	9
= 3.2 = G-E' interactions (specificity)	10
= 3.3 = A-A' interactions (stability and specificity)	15
= 3.6 = Anti-parallel leucine zippers	18
= 3.7 = LZ network design	20
= 3.9 = Conclusion	20
<b>= 4 = Folding</b>	<b>21</b>
= 4.1 = Overview	22
= 4.2 = Folding models	24
= 4.3 = Folding intermediates	27
= 4.4 = Diffusion-Collision-Desolvation (DCD) model	31
= 4.5 = Summary	34
= 4.6 = Conclusion	36
<b>= 5 = Functional diversity</b>	<b>36</b>
= 5.1 = Transcription factors – bZIP, bHLH-LZ, HD-ZIP	37
= 5.2 = Immune response signalling – NF-kappaB pathway	39
= 5.3 = More kinases – PKG, ZIPK, DAPK	43
= 5.4 = Ion channels – AKAP	43
= 5.5 = Transport vesicles – SNARE	44
= 5.6 = Viral envelopes and capsids	44
= 5.7 = Innate antiviral defense – interferon induced Mx proteins	45
<b>= 6 = LZ in protein engineering</b>	<b>46</b>
<b>= 7 = Conclusions and Outlook</b>	<b>46</b>
<b>= 8 = References</b>	<b>48</b>

## **Abstract**

In this essay we attempt to reconsider the concept of the “Leucine Zipper” (LZ) protein oligomerization motif. Reasoning on the wealth of existing data, we suggest that despite of the structural similarity with highly stable extended “Coiled Coil” motifs, on the functional level short and moderately stable “Leucine Zippers” might stand out as a distinct group. This family of oligomerization motifs facilitates highly specific combinatorial protein-protein recognition in the course of signal transduction events, thus going beyond the structural role of the extended “Coiled Coils”. Summarizing existing empirical knowledge on the stability and specificity of LZ we demonstrate how a simple set of rules, applied in the context of the universal coiled coil scaffold, creates a robust LZ interaction vocabulary. Being a highly abundant protein oligomerization motif, Leucine Zippers might account for coupling of distinct protein signalling pathways into a unified intracellular signalling network. In the last part of this essay we provide examples demonstrating prevalence of the LZ-mediated signal transduction and illustrate applicability of “LZ code” formalism to interpret evidences of couplings between cytoplasmic and nuclear signalling networks.

## = 1 = Introduction

Decryption of protein one-dimensional sequence from the corresponding nucleic acid sequence provided one of the key advancements for the emergence of genetic engineering and molecular biology. Unfortunately, high complexity of protein three-dimensional structures defers the widespread advent of the protein engineering. Namely, decryption of protein 3D structure from its primary sequence is not accessible yet and remains one of the fundamental frontiers in modern biology, generally referred to as the “protein folding” problem. One of the main motivations for solving this problem is the desire to understand and accurately predict interactions between proteins and other biomolecules within the cells. This knowledge is vital for understanding of a wide range of cellular processes governed by protein signal transduction (for example transmittance of extracellular signals to the transcription machinery). As a rule, these interactions are defined by extended and often highly dynamic 3D protein interfaces, making *ab initio* prediction of these interactions a daunting task, which cannot be solved at the current state of science and technology. However, a small part of this problem appears solvable already today. Leucine Zippers (LZ) represent a family of abundant protein-protein interaction motifs. Being based on the well characterized coiled coil scaffold, Leucine Zippers allow to reduce the interaction prediction problem to a simple comparison of two linear LZ amino acid sequences. This does not bring us closer to solving the general “protein folding” problem, but omnipresence of Leucine Zipper-based protein interactions makes such “LZ code” formalism a useful tool for evaluation of protein interactions among plethora of LZ-mediated signalling pathways.

Leucine zippers belong to the class of coiled coil structural motifs, arguably the simplest and the most ubiquitous mediators of protein-protein interactions (1, 2). The members of the LZ class exhibit extreme thermodynamic stability owing to the prevalence of leucine residues at the key positions of their hydrophobic interface. This allows reduction of a minimal peptide length required for oligomerization to three (3), sometimes even two (4, 5) heptad repeats. Based on this high stability per heptad the LZ motifs and fragments were proposed to serve as folding triggering sequences in the context of extended coiled coil structures (6, 7).

Based on the data from genome sequencing projects, coiled coils are established as the most abundant protein motif and are predicted to be found in 5-10% of all proteins (1). Their importance and versatility both *in vivo* and *in vitro* is underscored by the amount of literature

available on the topic, with a substantial number of valuable reviews appearing in the recent years (2, 8-11).

Contrary to the “elder” members of the coiled coil class of proteins, which are “obligatory oligomers” and mainly participate as structural cores in macromolecular ensembles (filaments, extracellular matrices, cytoskeletal networks, spacers, stalks, etc), LZ motifs represent “transient oligomers”, predominantly found in the signalling and regulatory proteins (receptors, kinases, transcription factors, etc), which reflects the transient nature of these interactions.

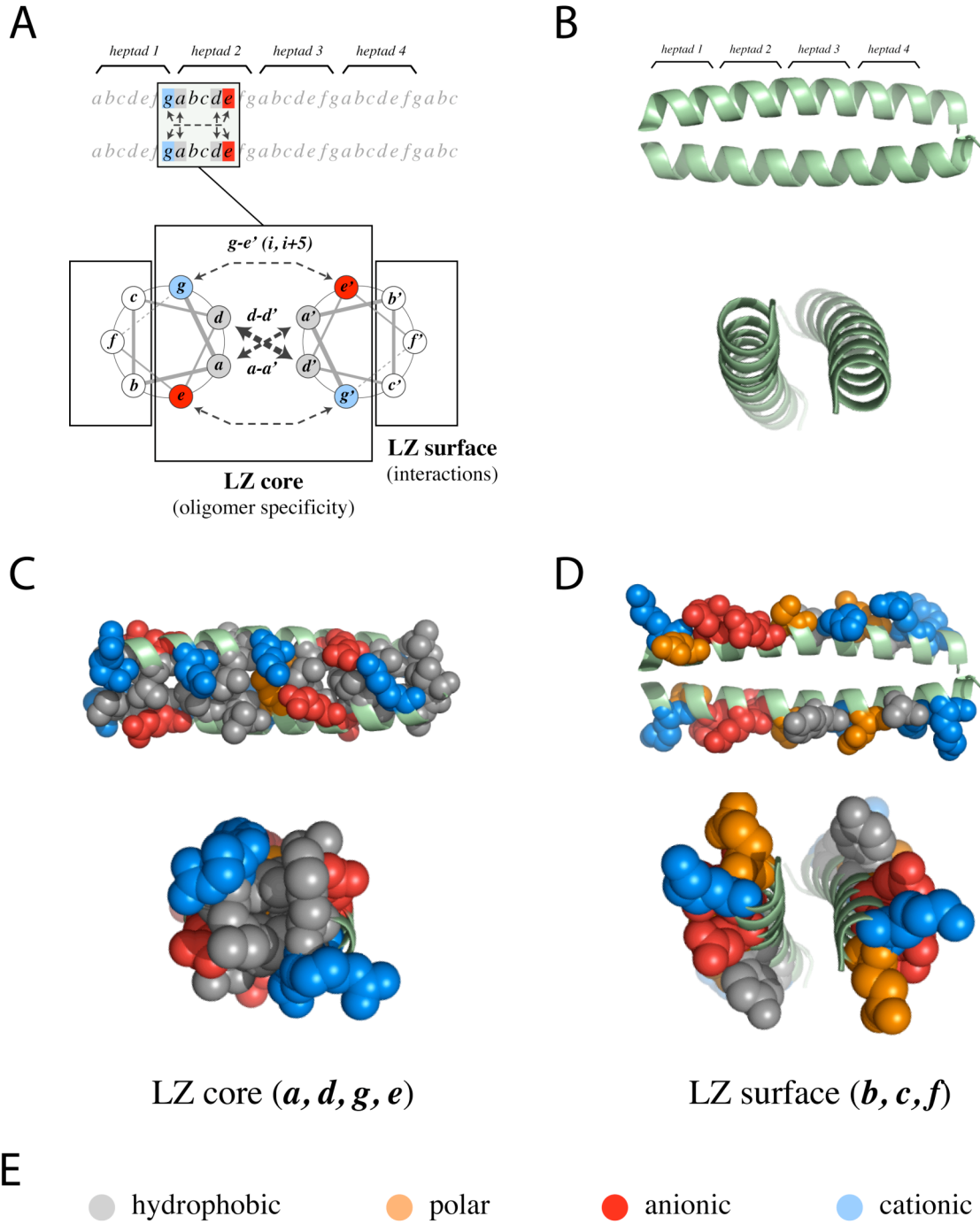
The Leucine Zipper motif was originally discovered in 1988 in the family of transcription factors named bZIP (basic region leucine zippers) (12). Shortly after its discovery, their presence was revealed in a much broader array of proteins (13, 14). During the two past decades the LZ motif has been actively employed as a model for protein folding (15, 16) and protein engineering studies (17, 18) (and references therein).

Here we review the existing data on the structure, interaction specificity and folding characteristics of LZ motifs, revealing the molecular mechanisms underlying LZ-enabled protein signalling. We discuss the omnipresence of LZ motifs and illustrate their ability to couple distinct protein signalling pathways.

## **= 2 = Structure**

### **= 2.1 = Primary – heptad repeat**

Primary structure of leucine zippers, as coiled coils class of proteins, is defined by characteristic seven residue (heptad) sequence repeat –  $(a\ b\ c\ d\ e\ f\ g)_n$ , where the pattern is formed by hydrophobic residues at the *a* and *d* positions, charged residues at the *e* and *g* positions, and generally polar residues elsewhere (19) (Figure 2.1).

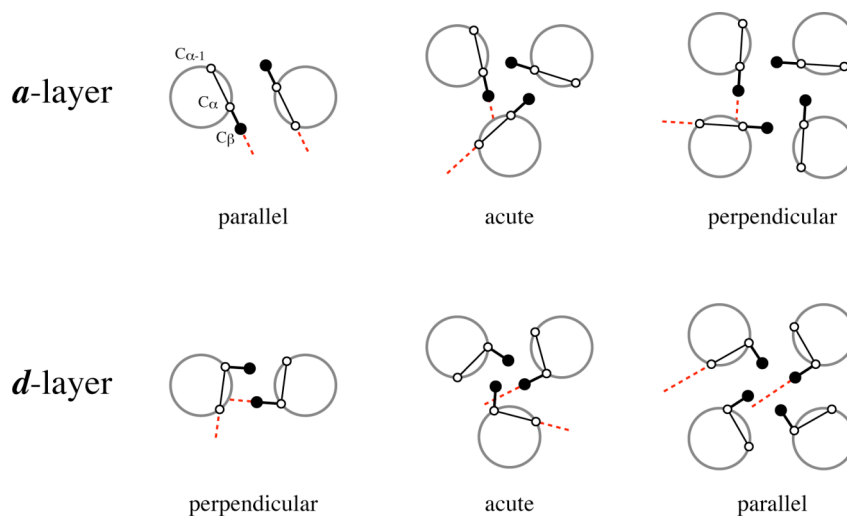


**Figure 2.1.** LZ structure and interactions. (A) Linear and wheel representation of coiled coil heptad repeat structure. (C) LZ core formed by hydrophobic  $d-d'$ ,  $a-a'$  and electrostatic  $g-e'$  interactions. (D) LZ surface  $b, c$  and  $f$  positions generally do not affect stability and specificity of LZ structure.

## = 2.2 = Secondary and tertiary – stability and stoichiometry

Regular amphiphatic primary sequence drives polypeptide assembly into a supercoiled structure, with the knobs-into-holes packing of hydrophobic  $a$  and  $d$  side chains at the interacting interface (20, 21). Charged residues at the  $e$  and  $g$  positions pack over the hydrophobic core effectively shielding it from the solvent, stabilizing the structure by inter-

chain *g-e'* electrostatic interactions and providing essential determinants for specificity of dimerization interface (22, 23) (more details below).



**Figure 2.2.** Packing interactions in the coiled coil hydrophobic core (see text for details).

The key structural difference of leucine zippers from other coiled coils is almost exclusive presence of leucine residues in the *d* positions of the hydrophobic core (12), which essentially defines their dimeric nature. As shown by Pehr Harbury and colleagues (24) stoichiometry of a coiled coil is mainly determined by side chain packing geometry of the hydrophobic residues in the *a* and *d* positions of the interface, which varies systematically between different oligomeric states (reviewed in (2)). Briefly – packing topology of coiled coil hydrophobic core is distinguished by the orientation of  $C\alpha-C\beta$  bond of the hydrophobic residues (*a* and *d* positions) relative to the peptide bond of the opposing helix (Figure 2.2). In **parallel** orientation the  $C\alpha-C\beta$  vector projects out of the dimer interior allowing more space between residues and thus favoring  $\beta$ -branched side chains (Ile, Val, Thr), where methyls branching from  $C\beta$  project back into the core, providing efficient Van der Waals interactions. Conversely, in **perpendicular** orientation  $C\alpha-C\beta$  vector projects directly into the core, limiting space available for the sidechains branched at  $C\beta$ , simultaneously providing excellent packing space for  $C\gamma$ -branched Leucines. Knobs-into-holes folding topology of the dimeric coiled coils brings residues of the *a*-layer into parallel orientation, and *d*-layer – into perpendicular. Thus, sequences bearing Leucines in *d* positions, and beta-branched residues in *a*, are very likely to fold into dimers. The situation is reversed in the tetrameric coiled coils fold: *a*-layer adopts perpendicular orientation, and *d* – parallel. Therefore this fold is favored by the sequences containing (Ile, Val, Thr) in *d* positions, and Leu - in the *a*. Topology of trimeric coiled coil fold is less restrictive - it has an intermediate (“acute”) geometry in both *a* and *d* layers - thus allowing more versatile sequence patterns.

### = 2.3 = Quaternary – specificity

Regular topology of interactions within the coiled coil motif together with a diverse set of available amino acid side-chains, provides LZ with a wide range of stabilities and specificities, allowing them to form both homodimeric and heterodimeric structures depending on the motif composition. Moreover, a significant fraction of natural LZ motifs exhibits a wide range of intrinsic specificity allowing them to form a variable set of heterodimeric pairs. This variability of specificities is a fundamental property that enables the LZ transcription factors to assemble combinatorial regulatory networks based on their LZ motifs. These networks - bZIP, bHLH-LZ, HD-ZIP - are amongst the most advanced regulatory networks developed by eukaryotic species (25), and have evolved as key regulators in many processes, ranging from cell metabolism to tissue differentiation (26). The rules governing interaction specificity within these networks have been thoroughly characterized during last two decades, and are mainly defined by electrostatics of ***g-e'*** couplings and polar interactions of the ***a-a'*** pairs, as discussed in more details below.

### = 3 = Stability and specificity

Core packing at ***a*** and ***d*** positions, together with ionic interactions between ***e*** and ***g*** positions are the key factors influencing stability and specificity of the coiled coil assembly. Applying reductionist approach to the most widely studied family of LZ proteins – bZIP TFs, three main interactions can be distinguished for the analysis of thermodynamic contributions to stability and specificity of the LZ interface (Figure 2.1):

- 1) ***d-d'*** interactions (primarily hydrophobic & VdW > defining stability)
- 2) ***g-e'*** interactions (primarily electrostatic & VdW > defining specificity)
- 3) ***a-a'*** interactions (mixed hydrophobic/VdW/electrostatic > defining stability and specificity)

Most of currently existing data on the weights of these contributions to the stability and specificity of leucine zipper motifs was produced by Charles Vinson group through application of double-mutant thermodynamic cycle analysis (27) in the context of LZ motif from bZIP factor VBP for ***d-d'*** (28), ***g-e'*** (23, 29) and ***a-a'*** (7, 30) pairs. Obtained results were largely corroborated by studies performed by Robert Hodges group (31-33), who targeted predominantly homodimeric interactions in the context of engineered coiled coils stabilized by covalent cross-linking. However, highly convoluted oligomerization equilibrium exhibited by engineered peptides in the latter cases, in the absence of high-



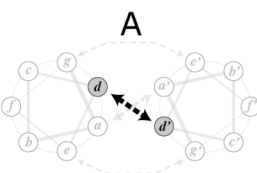
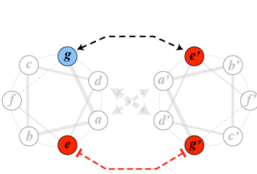
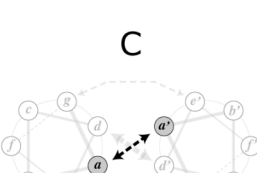
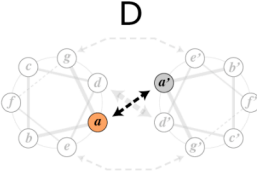
resolution structural data and double-mutant cycle free energy analysis urges to treat these data with caution when applied to canonical LZ motifs.

Detailed review of bZIP LZ stability and specificity, as well as specificity-based classification of bZIP transcription factors can be found elsewhere (10). Herein we provide a general summary on the topic, along with some contextual re-evaluation of available data.

### **= 3.1 = D-D' interactions (stability)**

Hydrophobic *d-d'* interactions are the key stabilizing component and the distinctive feature of the LZ family. Efficient packing of Leucine side chains in the *d* positions of the knobs-into-holes topology dramatically stabilizes the dimeric coiled coil interface (28), to a large extent defining the stoichiometry of the complex (24). Importantly, stability is conferred not only by the hydrophobic effect (burial of the hydrophobic side-chain in the protein interior, shielding it from the polar solvent) but also through Van der Waals interactions (efficient packing of the sidechain against neighboring residues). The latter contribution provides leucine with up to 5.2 – 5.9 kcal/mol/pair (contribution from one heptad) advantage in packing energy over similarly sized methionine and isoleucine pairs (28) (Table 3.1-A). 3D structure modeling suggests that the favorable rotamer conformations of beta-branched Ile and Val side-chains produce interhelical clashes between the C $\gamma$ 2 methyls if placed into the *d*-position (28). Thus, energy required to compensate for the thermodynamically unfavorable rotamer conformation may account for a part of the remarkable stability difference between leucine and beta-branched residues. This stability compromise does not play a significant role in the case of long structural coiled coil proteins, where a variety of hydrophobic amino acids have been shown to occupy the *d* position of the amphipathic helix (34). However, stabilizing effect of the leucine side chain appears crucial for the short leucine zipper sequences involved in signal transduction, thus yielding near invariance of this residue in the *d* position of the interface (28, 33).

Analyzed solely in the context of bZIP motifs, the role of *d*-position in determining the LZ interface specificity is apparently underestimated. For example in the Myc/Max/Mxd family of bHLH-LZ transcription factors, *d*-position histidine of Max protein forms a unique buried salt bridge with anionic sidechains in the heterodimerization partners, which defines the specificity of this network (35, 36). Thus, it is important to recognize that empirical dimerization rules discussed here provide only a part of the “LZ code” definition.

	Position	Residue	$\Delta\Delta G_{A-A}$ Energy difference [kcal/mol/pair]	$\Delta\Delta\Delta G_{int}$ Coupling energy [kcal/mol/pair]	Stability	Side chain / interaction type	Oligomer specificity	
<b>A</b> 	<b><i>d-d'</i></b>	L-L	-9.2		+	aliphatic		
		M-M	-4.0					
		I-I	-3.3					
		V-V	-2.2					
		C-C	-1.9					
		A-A	0.0					
		S-S	+0.5					
<b>B</b> 	<b><i>g-e'</i></b>	R-E	-1.55	-1.07	+	oppositely charged	(-) homodimers	
		K-E	-1.42	-0.91				
		E-R	-1.30	-0.45				
		E-K	-0.95	-0.25				
		Q-Q	-1.17	-0.03	-	~ identically charged		
		Q-K	-0.83	0.11				
		E-Q	-0.73	0.17				
		Q-E	-0.46	0.20				
		K-Q	-0.71	0.28				
		Q-R	-0.79	0.30				
		R-Q	-0.58	0.38				
		K-K	-0.34	0.45				
		K-R	-0.32	0.62				
		R-K	-0.10	0.66				
		E-E	+0.38	0.80				
		R-R	-0.10	0.81				
<b>C</b> 	<b><i>a-a'</i></b> (symmetric / homodimer)	I-I	-9.2	-0.9	+	aliphatic	(+) homodimers	
		V-V	-5.4	-0.7				
		L-L	-5.2	-0.6				
		N-N	-2.4	-0.5	~	polar		
		T-T	+0.8	+0.2				
		S-S	+1.8	+0.2				
		K-K	+2.9	+0.3	-	charged	(-) homodimers	
		R-R	+4.6	+1.2				
		<b>E-E</b>	<b>+6.0</b>	<b>+2.1</b>				
<b>D</b> 	<b><i>a-a'</i></b> (asymmetric / heterodimer)	(N,T,S) • (K,R,E)	+1.8	+0.1	~	charged • polar		
		(K,R,E) • (I,V,L)	-1.1	+0.2		charged • aliphatic		
		(I,V,L) • (I,V,L)	+5.5	+0.4		aliphatic • aliphatic		
		(K,R,E) • (K,R,E)	+3.9	+0.6	-	charged • charged		
		(N,T,S) • (N,T,S)	+1.1	+0.9		polar • polar		
		<b>(I,V,L) • (N,T,S)</b>	<b>-0.6</b>	<b>+2.3</b>		<b>polar • aliphatic</b>		
					--			

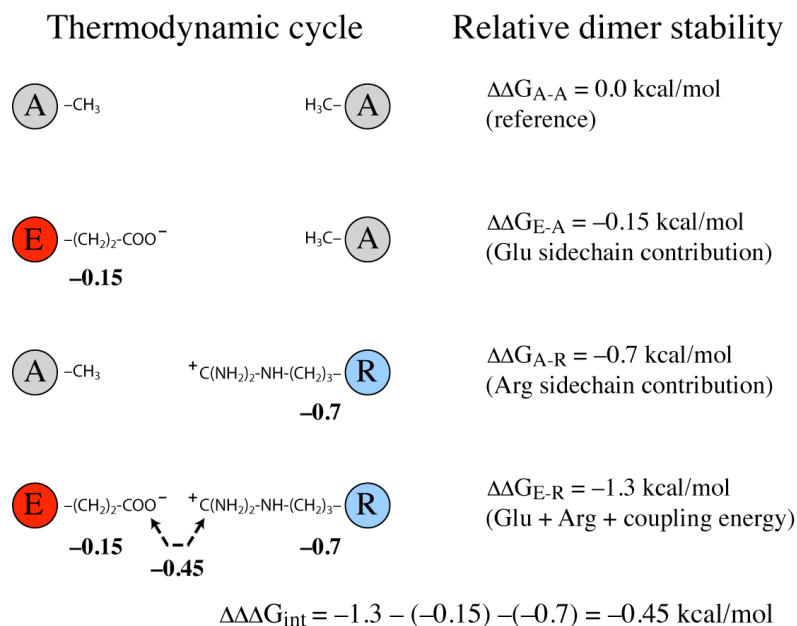
**Table 3.1.** Free energy differences ( $\Delta\Delta G_{A-A}$  [kcal/mol/pair] – useful to compare between LZ interaction types) and coupling energies ( $\Delta\Delta\Delta G_{int}$  [kcal/mol/pair] – useful when comparing pairs within one LZ interaction type) of common LZ coupling relative to a pair of alanines. Values obtained from LZ dimer thermal stabilities in 12 mM  $PO_4$ , 150 mM KCl, pH 7.4. Data reproduced from (A) *d-d'* (28), (B) *g-e'* (23, 29), (C) and (D) *a-a'* (7). For *g-e'* and *a-a'* interactions individual pairs are sorted according to the coupling energy strengths, and grouped in four categories:  $\pm 0.2$  kcal/mol (neutral),  $\leq 0.2$  kcal/mol (stabilizing),  $\geq 0.2$  kcal/mol (destabilizing),  $\geq 2$  kcal/mol (strongly destabilizing). Free and coupling energies for heterodimeric *a-a'* interactions (D) are averaged according to the residue type; full set of energies can be found in Table 3.2.

### = 3.2 = G-E' interactions (specificity)

**G-E'** interactions primarily involve charged amino acids with long aliphatic side-chains (Arg, Lys, Glu, Gln) (22), which simultaneously brings electrostatic, VdW and hydrophobic effects into play.

Compared to a pair of alanines, the most common bZIP ***g-e'*** salt bridges stabilize the coiled coil dimer by 1.3-1.6 (ER-RE) and 1-1.4 (EK-KE) kcal/mol/pair (Table 3.1-B). Remarkably, even identically charged Arg-Arg and Lys-Lys ***g-e'*** pairs have stabilizing effect, contributing respectively 0.1 and 0.34 kcal/mol/pair more energy than a pair of alanines. These repulsive electrostatic interactions are considered to be largely compensated by increased hydrophobic burial and favorable VdW interactions between the methylenes of ***g/e*** sidechains and hydrophobic core of the structure (21, 29, 35, 37, 38). Compared to alanine, the only destabilizing effect is shown by a pair of glutamates, which reduces the dimer stability by 0.38 kcal/mol/pair. Obviously two methylenes of a glutamate have less compensatory effect than three methylenes of an arginine and four methylenes of a lysine, with net energy differences markedly conforming ~0.5–1 kcal/mol protein stability gain commonly observed upon burial of additional methylene (39).

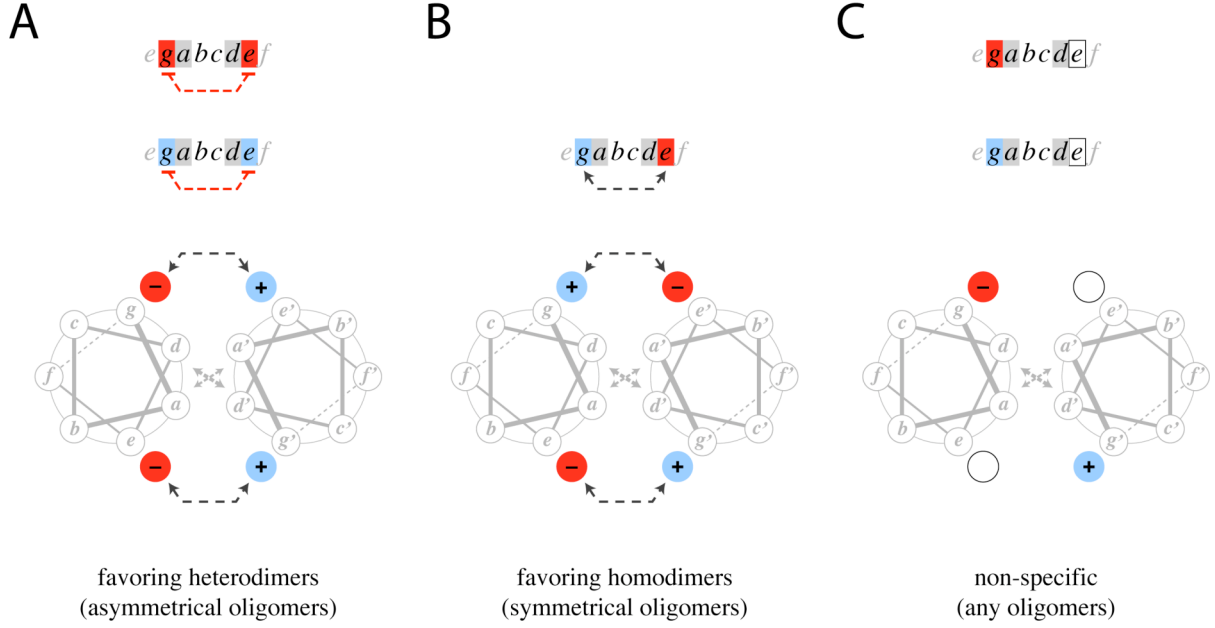
The overall contribution of interhelical salt bridges to the stability of leucine zippers for a long time has been a matter of debate (23, 29, 38, 40-42). The issue has been recently resolved by Hans Rudolf Bosshard and Daniel Marti, showing that the net thermodynamic contribution of a salt bridge is balanced between favorable charge-charge interaction, unfavorable desolvation energy and background interactions (such as coupling with the dipole moment of the helix) (43, 44). As it is evident from the Table 3.1, the effect of ionic ***g-e'*** couplings compared to hydrophobic core is rather moderate, and in the context of a canonical LZ heptade will be offset by energies of ***a-a'*** and ***d-d'*** couplings. Nevertheless, as will be discussed in the next section, the ionic interactions have a potential to regulate specificity of oligomerization by modulating kinetics of early steps of LZ folding process, when ***a-a'*** and ***d-d'*** interactions have not yet stabilized the structure. In this arrangement the long-range Coulombic interactions between charged side-chains shall be able to determine the specificity of coiled coil formation. The magnitude of these interactions for particular pairs of sidechains is most efficiently evaluated employing the concept of coupling energy, which is defined as the energy conveyed by the mutual interaction of two residues, devoid of energy contributions from isolated side-chains (23, 27) (Figure 3.1). For example coupling energy of E-R pair ( $\Delta\Delta G_{\text{int}} = -0.45$  kcal/mol) can be estimated as total E-R contribution to the dimer stability ( $\Delta\Delta G_{\text{A-A}} = -1.3$  kcal/mol) devoid of the stability contributions of individual E ( $\Delta\Delta G_{\text{E-A}} = -0.15$  kcal/mol) and R ( $\Delta\Delta G_{\text{A-R}} = -0.7$  kcal/mol) side-chains ( $\Delta\Delta G_{\text{int}} = \Delta\Delta G_{\text{A-A}} - \Delta\Delta G_{\text{E-A}} - \Delta\Delta G_{\text{A-R}}$ ) (23).



**Figure 3.1.** Thermodynamic double-mutant cycle for the Glu-Arg interaction. Measurement of thermal stabilities of four dimers yields three energy differences relative to a pair of alanines. Coupling energy ( $\Delta\Delta\Delta G_{\text{int}}$ ) of Glu-Arg ionic interaction is obtained by subtracting individual contributions of Glu and Arg sidechains from overall stability of the dimer.

Employing this concept the *g-e'* interactions can be arranged on a more reliable thermodynamic scale, defined by pure coupling energies devoid of stabilities conferred via interactions with the core of the molecule (Table 1, column  $\Delta\Delta G_{\text{int}}$ ). On this scale the most stabilizing interhelical coupling energies, on the order of  $-1$  kcal/mol/pair, are shown by R-E and K-E pairs, while the most destabilizing, on the order of  $+0.8$  kcal/mol/pair, by repulsive E-E and R-R couplings (23, 29). Importantly, coupling energies do not cluster and instead are uniformly distributed over the accessible energy scale. This diverse range of attractive, neutral and repulsive couplings available within common coiled coil scaffold, multiplied by the number of variable positions (8 in an average 4-heptad LZ motif) creates an efficient combinatorial key-lock mechanism for definition of interaction specificity. Distribution of specificity determinants along the whole leucine zipper sequence allows regulation of populations of different dimers in accordance with their composition (i.e. dimers with more attractive interactions and fewer repulsive interactions would be favored over dimers with fewer attractive and more repulsive interactions). This gives a potential for establishing a complex signalling node, capable of emitting a rich output signal instead of a simple on/off event. Moreover, as highlighted by differences in reciprocal K-E/E-K ( $-0.91$  vs.  $-0.25$  kcal/mol) and R-E/E-R ( $-1.07$  vs.  $-0.45$  kcal/mol) pairs (29), coupling energies of *g-e'* interactions strongly depend on the context, extending the combinatorial nature of LZ interface even further. However this effect appears to step into place only when underlying a

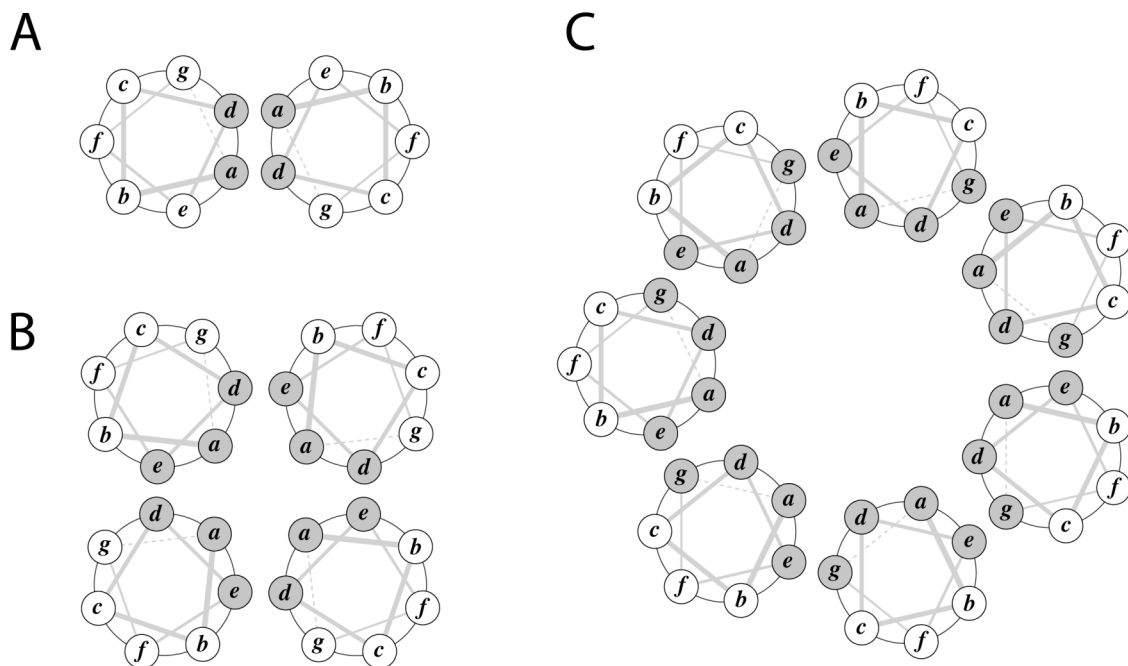
positions bear polar or charged side-chains, and is negligible in the case of purely aliphatic core (7).



**Figure 3.2.** Schematic representation of interhelical  $g-e'$  interactions in defining oligomerization specificity. (A) LZ with identically charged ( $i, i+5$ )  $g-e'$  residues – favoring heterodimerization, disfavoring homodimerization. (B) LZ with oppositely charged ( $i, i+5$ )  $g-e'$  residues – favor homodimerization. (C) LZ with non-ionic  $g-e'$  residues are not discriminative in oligomerization.

In the simplest case of homo- versus hetero-dimer formation, a pair of  $g-e'$  residues with the same charge (acidic + acidic or basic + basic) would favor asymmetric oligomerization – favoring heterodimers and disfavoring homodimers (Figure 3.2-A). A  $g-e'$  pair with alternating charges would favor symmetric oligomers (homodimers) and disfavor asymmetrical oligomers (heterodimers with mirrored charge allocation) (Figure 3.2-B). Non-charged side-chain would give the most liberal specificity range, allowing coupling with any type of residue (Figure 3.2-C).

In vivo these selective specificity mechanisms are successfully employed to decouple LZ-TF networks that operate in different functional realms. For example, specific  $g-e'$  electrostatic interactions define a subfamily of PAR factors, involved in regulation of circadian rhythms, precluding its cross-reactivity with other bZIP families (45). These considerations, together with the specificity rules conveyed by residues in  $a$ -positions, were successfully employed for classification of bZIP proteins based on their dimerization properties (46, 47).



**Figure 3.3.** Dependence of LZ oligomer stoichiometry on the size of continuous hydrophobic core. (A) Canonical LZ dimer with  $(a+d)$  hydrophobic interface. (B) Extended hydrophobic interface  $(a+d+e)$  yields a tetramer (23, 48). (C) Four-residue hydrophobic interface  $(a+d+e+g)$  yields up to a heptameric ensemble (18).

In addition to functional specificity (selection of dimerization partners),  $g-e'$  ionic interactions contribute to the structural specificity of LZ motifs, modulating register and orientation of monomer chains in the oligomeric ensemble (18, 49, 50). Furthermore, in the context of *in vitro* engineering studies,  $e$  and  $g$  positions can be employed for generation of high-order oligomers by extending the hydrophobic interface of the monomer chain. As originally shown by Harbury (24) (see “2.2 - secondary and tertiary structure” section above) – the stoichiometry of the coiled coil oligomers is to a large extent defined by the packing geometry of the residues occupying  $a$  and  $d$  positions of the sequence. However, a simpler rule might also be of some value in this respect – an estimate of continuous hydrophobic surface area carried by the coiled coil monomer. For example extension of a dimer-favoring 2-pair  $(a+d)$  hydrophobic interface (Figure 3.3-A), to a 3-pair  $(a+d+e)$  hydrophobic patches induces formation of tetramers (Figure 3.3-B) (23, 48), replacement of 14 sidechains in  $a$  and  $d$  positions with bulky tryptophan residues results in pentameric bundle (51), and extension of a 2-pair interface  $(a+d)$  to a 4-pair  $(a+d+g+e)$  creates high-order oligomers (52) with a heptameric coiled coil being the most striking structurally characterized example (Figure 3.3-C)(18).

### **= 3.3 = A-A' interactions (stability and specificity)**

The nature of this interaction has the most complex effect on the stability and specificity of LZ interfaces. Similarly to the residues in *d*-positions, packing of aliphatic side chains in *a*-position affects the stability and stoichiometry of the complex, with prevalence of C $\beta$ -branched amino acids (Ile, Val) (47) strongly favoring the dimeric structure of leucine zippers (24).

Similarly to Leucine in *d*-positions, isoleucine exhibits uniquely efficient side-chain packing in *a*-position, providing 9.2 kcal/mol/pair more energy than homotypic Ala interaction, and ~4 kcal/mol/pair more energy compared to similarly sized Leu or Val sidechains (30). However, as opposed by the extreme conservation of leucines in *d*-positions of the interface, isoleucine is a relatively infrequent residue in the *a*-position, with its occurrence probability being twice less compared to that of either leucine, valine and even asparagine (7). Sidechain selection working against the interface stability can be explained by two evolutionary advantages. First, as will be discussed below, incorporation of destabilizing polar residues provides additional mechanism for control over transcription factor functional (defining appropriate partners) and structural (defining stoichiometry and orientation) specificities. Thus, high occurrence of asparagine in the *a*-positions of bZIP factors highlights specificity-driven rather than stability-driven evolutionary pressures acting on these motifs. Secondly, moderate stability of the interface defined by high abundance of leucine and valine sidechains in the *a*-positions, as discussed in more detail in the “folding” section, reduces the activation energy needed for LZ dissociation, decreasing lifetime of the folded coiled coil state and elevating sensitivity of the LZ network to changes in external stimuli. This aspect underscores the notion of leucine zippers being a transient motif for signal transduction, rather than a static structural motif, as in the case of extended coiled coils.

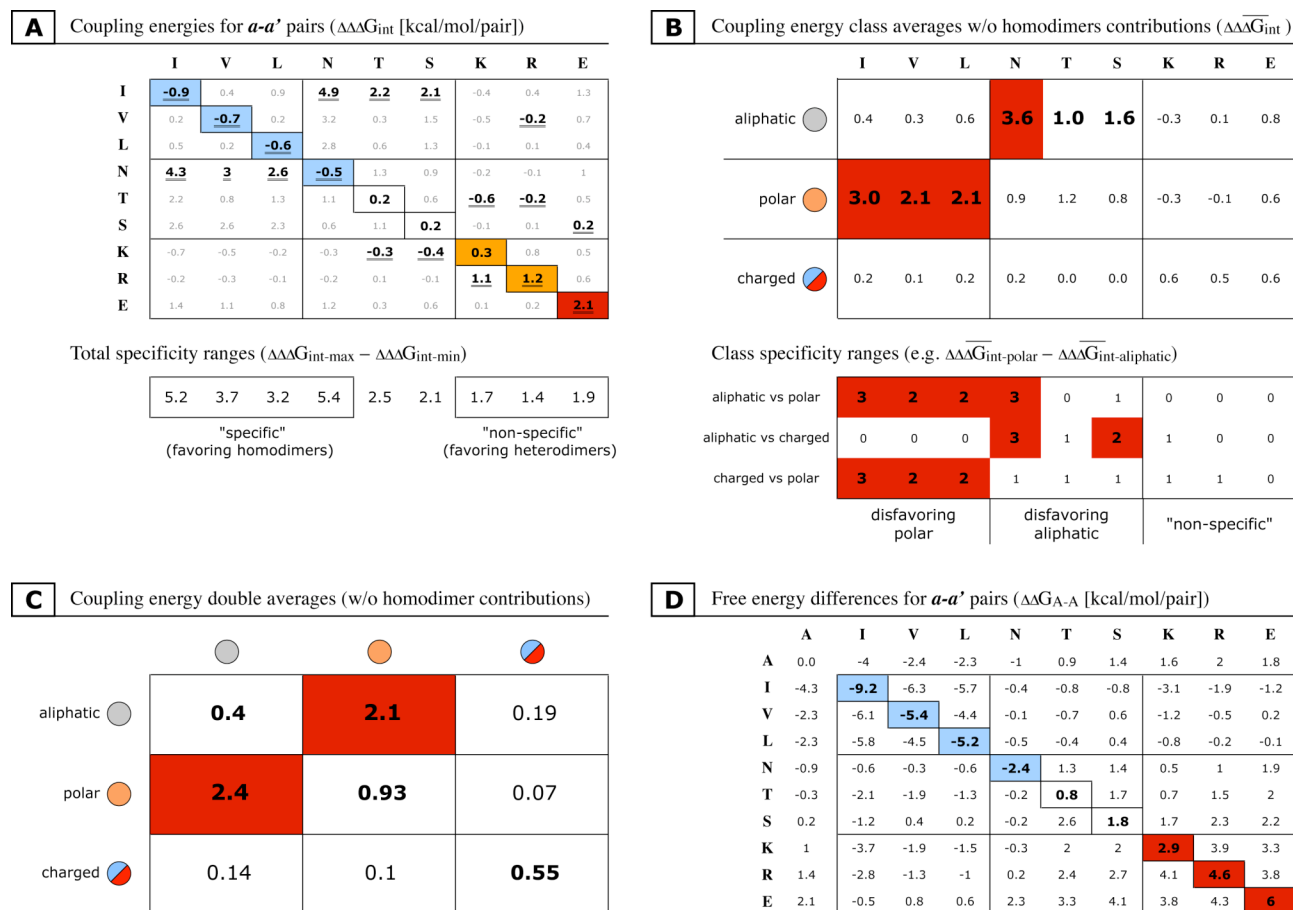
#### ***A-A' stability scale***

In addition to the “default” set of hydrophobic side chains, LZ factors often bear polar and charged residues in the *a*-positions of the interface. This creates an additional mechanism for control of specificity directing a wide range of homo- and hetero-dimerization events (7, 46, 47). Thermodynamic contribution of different residues to homodimeric *a-a'* interactions varies between stabilizing aliphatic, neutral polar and destabilizing charged sidechains (Table 3.1-C and diagonal in Table 3.2-D). This energy scale, relative to a pair of alanines, spans from –9.2 kcal/mol/pair for isoleucine to +6 kcal/mol/pair for glutamate (–0.9 kcal/mol/pair and +2.1 kcal/mol/pair in terms of coupling energies – Table 3.2-A), which signifies importance of individual *a-a'* couplings to the overall stability of the interface. Thus a vast 15 kcal/mol energy range is employed in regulation of LZ homodimerization specificity.



Similarly, a ~11 kcal/mol range of stability contributions is available for heterodimeric *a-a'* interactions (Table 3.1-D), facilitating control over heterodimerization specificity.

Interestingly, except interactions involving lysine sidechains, heterotypic interactions are predominantly destabilizing (Tables 3.1-D, 3.2-A&B).



**Table 3.2.** Specificity ranges of individual amino acids and amino acid classes in *a*-positions of the LZ interface. (A) and (D) data adapted from (7). As in Table 3.1 stabilizing coupling energies are highlighted blue, destabilizing – orange, and strongly destabilizing – red. (A) Coupling energies ( $\Delta\Delta G_{\text{int}}$  [kcal/mol/pair]) with corresponding specificity ranges defined by the difference between highest and lowest  $\Delta\Delta G_{\text{int}}$  values for particular residue. (B) and (C) Averaged coupling energies for heterodimeric couplings (i.e. devoid of homodimeric contributions) between different types of residues. (D) Free energy differences ( $\Delta\Delta G_{A-A}$  [kcal/mol/pair]) relative to the pair of alanines.

In the case of homotypic interactions, notable outliers are asparagine and lysine. Increased stability of polar Asn is thought to be brought by its favorable self-complementing hydrogen bonding (53). Meanwhile repulsive electrostatic interactions of lysine sidechain, as in the case of *g-e'* interactions, are offset by favorable hydrophobic burial and efficient VdW packing of its aliphatic backbone (33).

Likewise, destabilizing effect of polar and charged sidechains placed into heterotypic aliphatic context inversely correlates with their net hydrophobicity (number of methylenes in



the sidechain) (Tables 3.1-D, 3.2-A&B): Lys [-CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>-] > Arg [-CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>-] > Glu [-CH<sub>2</sub>CH<sub>2</sub>-] ≈ Thr [-CH<sub>2</sub>CH<sub>2</sub>-] > Ser [-CH<sub>2</sub>-] > Asn [-CH<sub>2</sub>-]. The highest destabilization effect is shown by asparagine, and similarly to its homotypic stabilizing effect is likely a consequence of an uncompensated hydrogen bonding (53). Unique properties of buried Asn sidechains for dimerization specificity control are underscored by its high abundance in naturally occurring LZ signalling networks (47). In addition to specificity control buried asparagines are known to be involved in control of LZ chain orientation (54, 55), register (9) and stoichiometry (24, 56), all factors possibly contributing to its frequent occurrence within LZ motifs.

Overall the *α-α'* stability scale (from most stable to most unstable):

- stabilizing: aliphatic (Ile > Val, Leu) and polar Asn
- neutral (hetero): charged • aliphatic, charged • polar
- neutral (homo): polar (Thr, Ser), charged/aliphatic (Lys)
- moderately destabilizing (hetero): polar • polar, charged • charged, aliphatic • aliphatic
- destabilizing (homo): charged Arg
- strongly destabilizing: charged Glu
- strongly destabilizing (hetero): polar • aliphatic

#### *A-A' specificity scale*

As suggested by Asha Acharya and coworkers (7) specificity of an individual amino acid in the *α*-positions can be estimated via the net coupling energy range they are capable to exhibit depending on the interacting sidechain. I.e. it is net energetic difference between the most stable ( $\Delta\Delta\Delta G_{\text{int-min}}$ ) and most unstable ( $\Delta\Delta\Delta G_{\text{int-max}}$ ) coupling exhibited by particular sidechain (Footer of Table 3.2-A). In the case if amino acid is highly selective (“specific”) it shall distinguish different pairing interactions, resulting in extended range of possible coupling energies. Conversely, non-selective (“unspecific”) residue shall not distinguish between different pairing sidechains, therefore its stability contribution shall not vary much depending on the partner.

On this scale isoleucine and asparagine show the greatest difference in coupling energies, indicating that those contribute the most to dimerization specificity (encourage homodimerization), while charged amino acids (K, R and E) show the least difference in coupling energies, suggesting that they contribute the least to dimerization specificity (i.e. tend to heterodimerize). Overall effect can be summarized as following - aliphatic residues (Ile, Val, and Leu) and Asn induce homotypic preferences in the LZ motif, polar Thr and Ser show neutral specificity, and charged sidechains (Lys, Arg and Glu) encourage heterodimerization (7).

To improve the precision of this analysis, we suggest to evaluate the abilities of individual sidechains to distinguish different classes of residues (i.e. aliphatic, polar, charged). For this purpose averages of heterodimeric coupling energies (devoid of homodimeric contributions) (Table 3.2-B) for particular residue classes shall be compared (Footer of Table 3.2-B). This allows, for example, to see that specificity (“specificity range”) of aliphatic sidechains is not uniform, and mainly relates to disfavoring polar partners, while being indifferent to aliphatic and charged sidechains. Furthermore, it becomes clear that polar serine and threonine also foster dimerization specificity similar to asparagine sidechain. Combined with Ser/Thr abundances in the natural LZ motifs (7), this observation points to their possible role as intermediate “specificity restrictors”, providing less stringent energy discrimination compared to the Asn sidechain.

Based on these revised “specificity ranges” the following conclusions for heterodimeric interactions can be made:

- (1) aliphatic residues strongly disfavor polar partners, but do not distinguish between other sidechain types.
- (2) correspondingly, polar residues strongly disfavor aliphatic partners, but are indifferent for other sidechain types.
- (3) charged sidechains do not differentiate between sidechain types.

These conclusions are most strikingly revealed upon further averaging of coupling energies within particular classes (Table 3.2-C). It is apparent that among heterotypic interactions the most unfavorable are those involving aliphatic and polar sidechains, while charged residues provide most stable couplings independent of the context.

Considering default hydrophobicity of the LZ core, the *a-a'* position specificity scale can be reformulated as following, from favoring homodimers (“specific”) to favoring heterodimers (“unspecific”):

- polar (Asn > Ser > Thr) (favoring homodimers & disfavoring heterodimers)
- aliphatic (Ile, Val, Leu) (favoring homodimers)
- charged Lys (favoring heterodimers)
- charged (Glu, Arg) (favoring heterodimers & disfavoring homodimers)

### **= 3.6 = Anti-parallel leucine zippers**

Along with the widespread parallel dimeric architectures, coiled coils are able to assemble complexes with an anti-parallel arrangement of helices. These structures seem to be poorly represented in nature, and therefore have not received due attention, although there seems to be an increase in interest to anti-parallel structures in the last years (9, 57). In principle these structures could fall under the same “leucine zipper” nomenclature, because of

the similar heptad repeat featuring conserved leucine side-chain at every seventh residue (a “*d*-position”) (9). However, similarly to the structural roles of extended parallel coiled coils, majority of the existing examples from the anti-parallel coiled coils are involved in formation of static structural cores, rather than dynamic signalling interfaces, therefore falling beyond the scope of this review. Nevertheless, a few key characteristics of these assemblies will be shortly highlighted below.

Similarly to their parallel relatives, antiparallel assemblies feature hydrophobic core formed by apolar side-chains in the *a* and *d* positions of the heptad repeat, stabilized by the electrostatic interactions between charged residues in the *g* and *e* positions. In the case of anti-parallel structures *a-a'* and *d-d'* hydrophobic interactions are replaced by *a-d'* and *d-a'* pairs, and *g-e'* electrostatic couplings are replaced with *g-g'* and *e-e'* pairs. As in the case of parallel structures most of the stability is conferred via the hydrophobic core, while specificity and anti-parallel chain orientation itself is mainly defined by Coulombic interactions between side-chains in the *g* and *e* positions (49, 50).

In addition, the potential of buried polar residues in determining structural integrity of anti-parallel coiled coils has been demonstrated by replacement of *a-d'* hydrophobic residues with a pair of asparagines (58). However, although *a-a'* polar interactions are an important specificity determinant for naturally occurring leucine zippers, the equivalent *a-d'* polar interactions has not been reported for anti-parallel assemblies.

Summing up – anti-parallel coiled coil interfaces seem to bear all the required determinants for assembly of signalling regulatory networks similar to those based on the leucine zipper interfaces. However one crucial difference creates an intrinsic limitation for anti-parallel coiled coil architecture within signalling cascades. This limitation stems from the packing efficiency of the hydrophobic core, which defines the structural integrity and stoichiometry of the coiled coil complex. As discussed above (see section 2.2 - secondary and tertiary structure), extreme stability and specificity of the parallel dimeric LZ interface is defined by very specific and efficient packing of hydrophobic side-chains within its core – *a*-layer side-chains adopt “parallel” orientation, while *d*-layer adopts a distinct “perpendicular” arrangement. Packing of *d*-position side-chains delivers most of the energy required to stabilize the interface, allowing certain flexibility at *a*-positions and thus providing a mechanism for control of stoichiometric specificity using polar residues in *a*-positions. In the case of anti-parallel structures the ability to differentiate between stability vs. specificity contributions is eliminated, since in these structures hydrophobic layers adopt a single geometrical type of side-chain arrangement, involving mixture of *a* and *d* side-chains (57). This lack of intrinsic structural specificity is demonstrated by heterogeneity of structural

species formed by 5-heptad coiled coil domain from hepatitis delta virus antigen (59) and structural instability of 10-heptad coiled coil from *E.coli* Seryl tRNA Synthetase (60).

### **= 3.7 = LZ network design**

Reviewing the discussed above LZ specificity rules, a few general remarks can be made. In the context of an isolated heptad homodimerization specificity can be achieved by incorporation of polar residues into ***a*** positions (moderately affecting homodimerization while strongly disfavoring heterodimerization) and incorporation of residues with alternating charges into the ***g-e'*** positions (since non-alternating ***g-e'*** charges disfavor homodimerization). Increased heterodimerization specificity can be achieved by incorporation of charged residues into ***a*** positions (disfavoring homodimeric while stabilizing most of heterodimeric couplings) and also introduction of identically charged residues in ***g-e'*** positions (seriously destabilizing homodimers).

Speaking about networks of factors, in the context of prevailing aliphatic side-chains in the ***a***-positions, the combinatorial specificity of a particular network can be increased by incorporation of polar residues (especially asparagines and serines) into the unique ***a***-positions of the interface – this will create a strong destabilizing effect for all except homotypic interactions (i.e. those having polar residues in the same position). Similarly, to couple a LZ-factor to a network defined by particular allocation of a buried polar sidechain, one has to place a polar side-chain in the corresponding location in the interface. To provide coupling between two networks specified by distinct allocations of buried polar residues, one shall incorporate charged residues in corresponding ***a***-positions of the interface (thus oligomerization within either of the networks will not involve unfavorable aliphatic • polar interactions). General increase in the amount of charged side-chains in ***a***-positions decreases the specificity and increases the range of interactions available for a particular LZ motif. Thus coupling of several specialized networks via a central hub requires more “unspecific” (destabilizing) residues in ***a*** positions of the heterodimerizing zipper, putting additional pressure on the optimization of its ***d-d'*** and ***g-e'*** interactions.

To selectively decouple distinct networks one has to increase the amount of specificity determinants – introduce polar residues in non-matching ***a***-positions and repulsive interactions between ***g-e'*** sidechains. These specificity determinants will not affect oligomerization within the family, while strongly disfavoring any interactions outside of it.

### **= 3.9 = Conclusion**

Clearly, for an adequate analysis of particular interface stability and specificity local context of described above interactions will play a very important role. For example

thermodynamic contribution of aliphatic side chains in the *d*-position varies up to 4 kcal/mol/pair (for leucine) depending on the neighboring residues in *a*-positions (28, 61); contribution of buried Asn residues is also context-dependent, varying on the order of 2 kcal/mol/pair depending on the environment (62); similar variability is shown by electrostatic *g-e'* couplings (discussed in (31)); polar and charged amino acids placed in *a*-positions energetically differentiate reciprocal orientations of overlying electrostatic *g-e'* pairs (7). As well it has been suggested that not only the sums of individual energies, but also the patterns of interactions define the stability and specificity of the LZ interfaces (47). Therefore, the issue of context still has to be resolved in more detail to increase the accuracy of predictions. Nevertheless, as verified by experimental data (63), even in the absence of more detailed contextual analysis, a simplified set of LZ specificity determinants already yields quite realistic predictions on oligomerization properties of canonical leucine zippers (46).

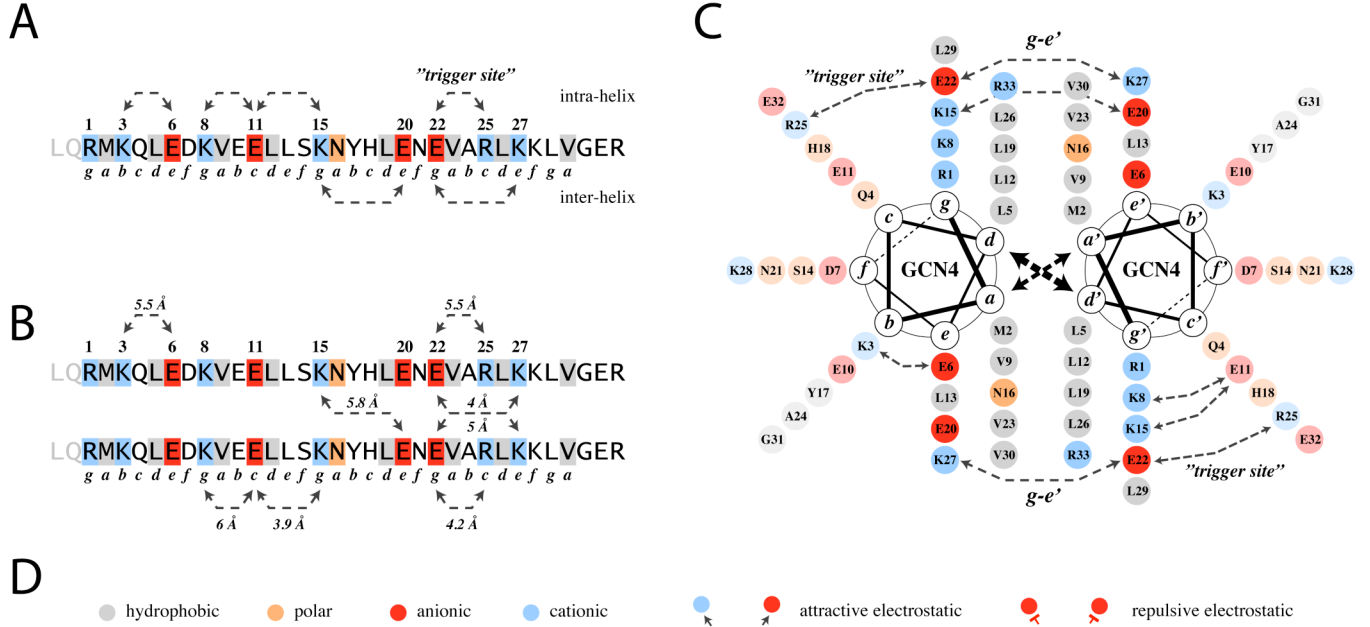
Importantly, beyond contextual dependencies, some gaps remain in the fundamental understanding of LZ specificity determinants. For example interactions within Myc/Max/Mxd network of oncogenic bHLH-LZ factors are specified by buried salt bridges involving *d*-position histidine on the Max side and *a*-position glutamate/aspartate residues on the Myc/Mxd side (35, 36). Another example refers to a group of plant bZIP TFs which employs a conserved proline residue in the *f*-position of the interface to restrict formation of homodimers, thus profoundly changing the topology of the signalling network (64).

Therefore it seems reasonable to apply described above simplified set of determinants only in the context of “canonical” LZ motifs, and only when comparing interactions with notable energetic differences, since subtle energy variances will be masked by the error imposed by these simplifications. Further advancements in our understanding of LZ interaction stability and specificity require more thorough sampling of the interaction space, and thus are expected to come from the systems biology approaches (7, 65).

## **= 4 = Folding**

The stability and specificity rules derived from thermodynamic properties of LZ interfaces provide only a partial insight into the nature of LZ-mediated signal transduction, showing the network equilibrium state at an infinite time limit. In addition to this “thermodynamic control”, protein signalling is highly dependent on the kinetics of particular interactions, including the presence of structured intermediates which provide specificity filters when signal transduction is coupled with folding process. These characteristics of protein folding landscape facilitate the “kinetic control” over signalling events, determining

the sensitivity of interaction to the changes in the “input signal” – i.e. the timescale of the signalling event. In the following section we review the existing knowledge on the folding of parallel dimeric leucine zipper motifs, to aid further understanding of LZ signalling mechanisms.



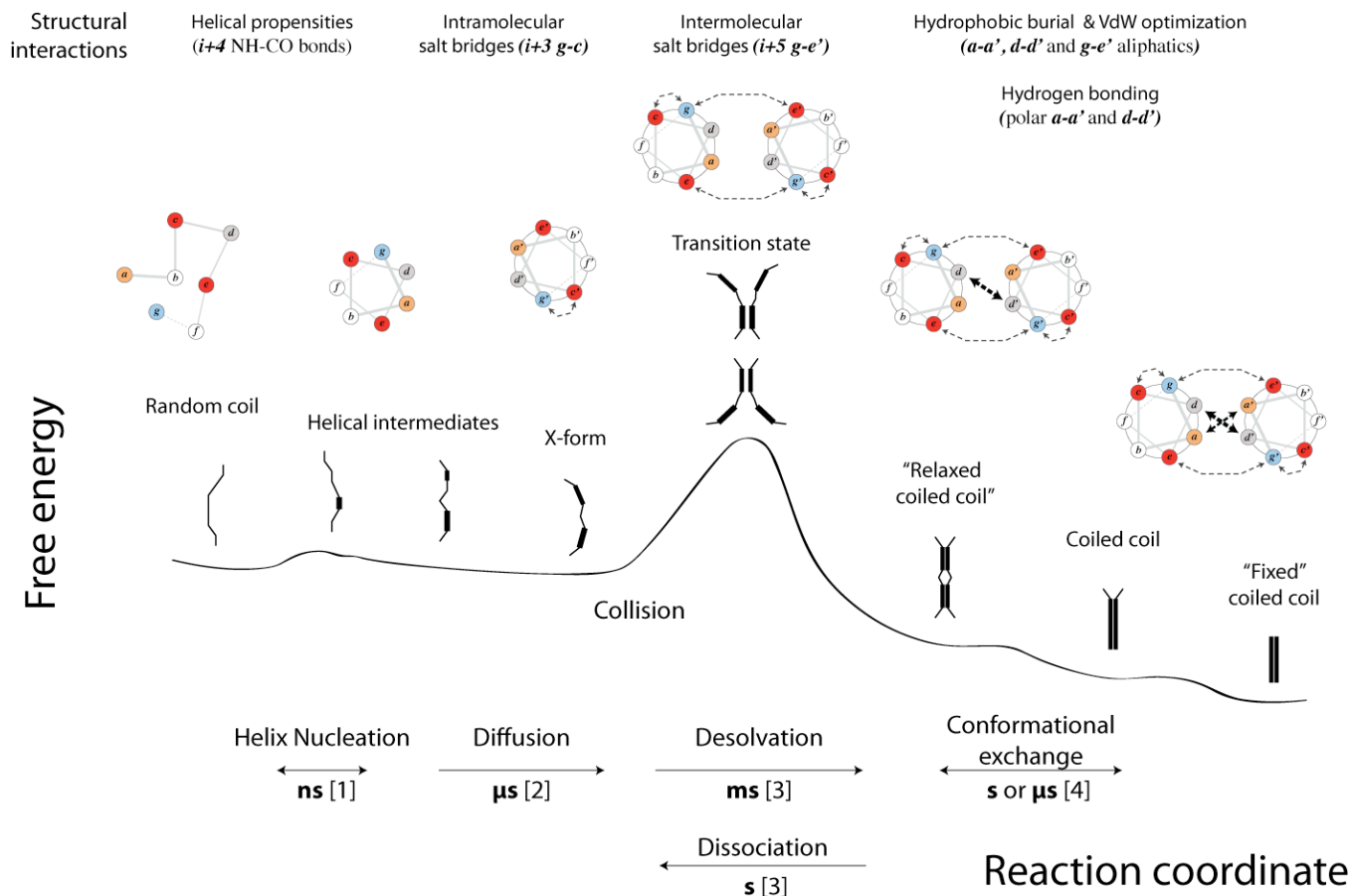
**Figure 4.1.** Electrostatic interactions within GCN4p1 (LZ-GCN4). Residue numbering according to GCN4p1 sequence. **(A)** One-chain linear notation, most useful for illustration of interactions within symmetrical homodimeric LZ motifs. **(B)** Two-chain linear notation, most useful for illustration of asymmetrical LZ motifs. **(C)** Helical wheel notation, useful in any situation. **(D)** Figure legend (same coloring is employed in all other figures with LZ motifs). Distances between charged atoms in **(B)** are based on the basic region + leucine zipper fragment of GCN4 bound to its consensus DNA sequence (pdb:1ysa).

## = 4.1 = Overview

For the most part our knowledge of LZ folding process is based on the studies of GCN4p1 – archetypical 33 amino acid peptide corresponding to the LZ motif of yeast transcription factor GCN4 (Figure 4.1). In addition a considerable amount of experimental data relates to engineered LZ motifs, designed to distinguish various contributions (hydrophobics (66, 67), electrostatics (68, 69), helical propensities (67, 70, 71)) to LZ folding landscape. For a long time the general view on LZ folding was that monomer chains are largely unstructured at the early stages of the folding process (67, 68, 72, 73), and that the main energy barrier in the folding direction is highly entropic by nature (i.e. determined largely by hydrophobic and VdW interactions of *a-a'* and *d-d'* couplings) (70, 74). However, later it became apparent that at least one helical intermediate is populated prior to the main

folding event (6, 66, 70, 75, 76). In addition to that, recent studies of Jun and Fos LZ motifs revealed significant contribution of enthalpic component (electrostatic and polar *g-e'* interactions) to the activation barrier in the folding direction (77, 78).

Here, rationalizing the available data, we propose that generalized folding process of short dimeric LZ motifs is best described by the Diffusion-Collision-Desolvation model (Figure 4.2). In this model, stretches of helical structure (corresponding to the “microdomain” elements of the original Diffusion-Collision model (79)) are primed by hydrogen bonds and stabilized by intra-helical salt bridges within LZ monomers at the early stages of the folding process (16, 76, 80). These intermediates collide in a diffusion-limited manner, with the probability of accessing productive transition state dependent both on the prominence of helical structures and the rate of collisions between these “microdomains”. In Diffusion-Collision-Desolvation model the main activation energy barrier is dependent on the long-range electrostatic interactions between the monomer chains – balanced between favorable “electrostatic guidance” (81) and unfavorable desolvation contributions (44, 82). These contributions are reflected in the enthalpic component of the free energy barrier (77) and provide an essential LZ specificity discrimination mechanism, based on the long-range Coulombic forces.



**Figure 4.2.** Diffusion-Collision-Desolvation model for LZ folding. References for kinetic rates:

[1] Helix nucleation, 1-17 ns (83-86).

[2] Theoretical diffusion-limited collision rate  $\sim 2.5 \mu\text{s}$  at 100  $\mu\text{M}$  peptide (66, 75, 87).

[3] Forward time constant (monomer lifetime) 0.7-25 ms at 100  $\mu\text{M}$  peptide; reverse time constant (dimer lifetime) 2-300 s (16, 67, 72, 74, 75, 80).

[4] Exchange time scale 0.4 s for GCN4-IzK analog (88); 0.2-1.2 s Jun-Fos analog (78);  $\sim 10 \mu\text{s}$  for crosslinked GCN4p1 (15).

Similar to the monomeric intermediates at the non-native side of the folding barrier, several groups have reported on existence of a stable dimeric intermediate at the native side of the folding barrier (15, 88, 89), designated here as the “relaxed coiled coil” state. The exact nature of this state has yet to be revealed, however repacking of the hydrophobic core within the central region of this structure (89), points to a possible rationale behind this transition. Namely, interactions involving polar buried residues in the  $\alpha$ -positions of the LZ interface (N16 in case of GCN4p1) were shown to manifest themselves only after the rate-limiting step in LZ folding process (90). Therefore, it seems plausible that the “relaxed coiled coil” state is defined by reorganization of VdW packing and hydrogen bonding established by buried polar sidechains. This reorganization decreases the stability of the final coiled coil state, reducing the height of unfolding activation barrier, thereby modulating the lifetime of the signalling event.

Combined DCD model explains LZ specificity discrimination mechanism based on the long-range electrostatic interactions between monomer chains, and elucidates essential “kinetic control” components on both sides of the main folding barrier. Combinatorial multiplicity of LZ interfaces discussed in the previous section, together with the flexible kinetic control of LZ folding landscape, provide the fundamental basis for the remarkable versatility of this motif in establishment of protein signalling pathways.

Summarized above aspects of LZ folding process are discussed in more detail below.

## **= 4.2 = Folding models**

### ***Two-state model***

For a number of years folding of LZ was considered a two-state process involving predominantly unstructured monomer and a fully-fledged coiled coil dimer (67, 68, 72, 73). In the two-state model folding starts as a collision of two unstructured monomer chains, followed by a “downhill” hydrophobic collapse resulting in formation of a folded coiled coil dimer. This does not mean an all-or-none synchronous structuring of the whole chain, but rather refers to the situation where all molecular conformations can be organized into two general groups divided by a single high-energy barrier. In the case of leucine zippers those



groups represent predominantly disordered monomers (M) and predominantly folded coiled coil dimers (CC):



In the two-state LZ folding approximation the transition state contains little if any secondary structure, and the highest energy barrier (rate limiting step) is primarily dictated by the diffusion processes:

$$k_f \propto D \quad (1b)$$

where D reflects the frequency of diffusion-limited collisions events.

As happened in the course of early LZ folding studies, depending on the sensitivity and time resolving capabilities of particular experimental setup, some non-two-state processes may appear as two-state because of short lifetimes and/or low stabilities of the folding intermediates.

### ***Diffusion-Collision model***

Eventually experimental data started to accumulate indicating that folding of LZ is better described by a diffusion-collision model (91), where at least one helical intermediate is populated prior to the main folding event (6, 66, 70, 75, 80, 92). As opposed to the two-state model, Diffusion-Collision theory relies on the existence of preformed structural elements, termed microdomains, which collide at diffusion-limited rates (79).



In this model the folding rate is dependent both on diffusion-mediated processes and “coalescence probability” term:

$$k_f \propto D \times \beta \quad (2b)$$

where beta (“coalescence probability”) corresponds to the fraction of collisions which are productive (leading to the transition state), embracing both the prominence of elementary microdomains (defined by  $k_1/k_{-1}$ ) and barriers mediating the coalescence step (e.g. probability of productive orientation at the moment of encounter – defined by  $k_2/k_{-2}$ ).

Most of researchers currently support the DC concept, agreeing that simple kinetic considerations strongly favor this model. Specifically, helix nucleation (i.e.  $M \rightarrow I^*$  transition) has been reported to occur on the nanosecond timescale (~1-17 ns) (83-86), while theoretical LZ monomer collision rate is 3 orders of magnitude slower (~2.5 microseconds at 100 μM peptide) (66, 75, 87)), and experimentally observed LZ folding rate is yet another 3-4 orders of magnitude slower than collision rate (0.7-24 milliseconds at 100 μM monomer concentration) (16, 67, 72, 74, 75, 80) (Table 4.1). Notable difference in timescales of individual folding steps (~ns helix nucleation → ~μs collision → ~ms dimer assembly),

indicates that: 1) at the moment of collision monomer chains contain a considerable amount of pre-formed helical structure, and 2) only a fraction of collisions leads to formation of the coiled coil dimer. Although it was shown plausible to design a coiled coil with negligible intrinsic helicity that would fold via pure two-state mechanism (69), natural occurrence of such monotonous sequences is unlikely and thus application of the collision-first two-state model would be an oversimplification.

Helix nucleation										
						$k_1$ (unimolecular rate constant)	time to helix nucleation (unimolecular)			
							$1/(k_1)$ [ms]			
						[M-1 s-1]				
Hummer, 2000							0.000001			
Thompson, 1997						1.E+08	0.000010			
Williams, 1996						6.E+07	0.000017			
Collision (theoretical)										
						$k_1$ (bimolecular collision rate constant)	time to collision @ 100μM			
							$1/(k_1*100\mu M)$ [ms]			
						[M-1 s-1]				
Durr, 1999						5.E+09	0.002			
Zitzewitz, 2000						3.E+09	0.003			
Holtzer, 2001						3.E+09	0.003			
Dimerization										
peptide	NaCl	buffer		temp	pH	$k_1$ (bimolecular association rate constant)	monomer lifetime @ 100μM (association time)	$k_{-1}$ (unimolecular dissociation rate constant)	dimer lifetime (dissociation time)	$K_d$
							$1/(k_1*100\mu M)$ [ms]		$1/k_{-1}$ [s]	
	[mM]	[mM]		[°C]		[M-1 s-1]		[s-1]		[nM]
Zitzewitz, 1995	GCN4p1	150	50	Phosphate	5	7	4.2E+05	24	0.0033	8
Zitzewitz, 2000	GCN4p1	150	50	Phosphate	15	7	2.2E+06	4.5	0.0183	8
Moran, 1999	GCN4p1	150-200	20	Acetate	10	5.5	2.0E+06	5.0	0.0068 <sup>†</sup>	3
Bosshard, 2001	BRLZ-GCN4 (62aa)	10	50	Phosphate	5	7.4	5.5E+06	1.8	0.01	2
	BRLZ-GCN4 (62aa)	10	50	Phosphate	30	7.4	1.4E+07	0.7	0.6	43
Ibarra-Molero, 2004	GCN4p1-M2V	150	50	Phosphate	15	7	2.2E+06	4.6	0.45	205
Steinmetz, 2007	GCN4p1	150	10	Phosphate	5	7.4	6.6E+05	15	0.17	256
AVERAGE:						4.E+06	8	0.18	88	
Nikolaev, 2007 (X-form)	Leu-Gln-GCN4p1	50	50	Acetate	37	3.2	8.6E+04	116	2.6	30000
Dimerization (cross-linked vs dimeric)										
peptide	NaCl	buffer		temp	pH	$k_1$ (association rate constant)	association time	$k_{-1}$ (dissociation rate constant)	dissociation time	
Moran, 1999	cross-linked GCN4p1	150-200	20	Acetate	10	5.5	7.5E+03	0.133 <sup>*</sup>		
	GCN4p1	150-200	20	Acetate	10	5.5	2.0E+06	5	0.0068 <sup>†</sup>	3
Wang, 2005	cross-linked GCN4p1	N/A	50	Phosphate	T-jump	7		0.100 <sup>*</sup>		
								0.010 <sup>*</sup>		

\* - unimolecular process

† - graphical estimate from Chevron plot

**Table 4.1.** Summary of LZ folding rates. Corresponding references (column 1): Hummer, 2000 (84); Thompson, 1997 (85); Williams, 1996 (86); Durr, 1999 (66); Zitzewitz, 2000 (75); Holtzer, 2001 (87); Zitzewitz, 1995 (72); Moran, 1999 (67); Bosshard, 2001 (74); Ibarra-Molero, 2004 (80); Steinmetz, 2007 (16); Nikolaev, 2007 (76); Wang, 2005 (15).

### **= 4.3 = Folding intermediates**

The transition from the two-state to the Diffusion-Collision folding model of LZ folding, was accompanied by discoveries of stable folding intermediates at both non-native (monomeric) and native (dimeric) sides of the folding barrier. Related findings are summarized below, and importance of both intermediate types in establishing “kinetic control” over LZ signalling process is discussed.

#### ***Triggering sequence***

According to the DC model, the protein folding rate increases proportionally with the endurance of preformed structural elements (helical segments in the case of coiled coils). Formation of these elements reduces conformational entropy of the peptide chain, thus increasing the population of association-competent conformational states and decreasing the activation energy required to access the transition state. The change from the two-state to DC view on LZ folding process was initially sparked by the emergence of the “triggering sequence” concept. This concept suggested that a conserved set of electrostatic interactions is present in a diverse set of coiled coil motifs, which induces formation of helical structures in the monomeric chains, increasing folding rates of short LZ sequences and providing a folding nucleation site for extended coiled coils (6). In archetypical GCN4p1 this “triggering sequence” is exemplified by the cluster of interactions around Glu22-Arg25 salt bridge (Figure 4.1). A diverse set of experimental studies has indeed confirmed the importance of the E22-R25 intramolecular salt bridge for helix stabilization in the early steps of the GCN4p1 folding process (16, 70, 75, 80).

Nonetheless, researchers later concurred that alpha-helix can be stabilized in many different ways, and that within extreme diversity of the coiled coil motifs a specific consensus “triggering sequence” is unlikely to exist. This conclusion is supported by experimental studies showing that early folding kinetics of LZ are also highly sensitive to perturbations in peptide intrinsic helicity (67, 70) and chain hydrophobicity (66, 67). Furthermore, as shown by Darin Lee and coworkers (71) the presence of the proposed triggering sequence *per se* does not guarantee successful folding of GCN4p1 analogs.

The final remark regarding the “triggering sequence” concept relates to the multiplicity of the accessible protein folding pathways, as proposed within original DC theory (79). Although not disallowing the existence of uniquely robust folding pathway, the DC model explicitly permits multiple folding routes to be attained (79). Individual folding pathways thus depend on the properties of particular microdomains, providing an error-resistant folding landscape in the case of multiple-microdomain architecture (i.e. the protein is less likely to be trapped in the local energetic minimum). As shown by Liam Moran (67), GCN4p1 sequence

appears to bear three helix-nucleating regions – approximated at the N-term, C-term and at the center of the sequence. This correlates with the reported heterogeneity of the GCN4p1 transition state (69, 75) and reflects the adaptive characteristics of GCN4p1 folding landscape. When a single-site mutation is made that disrupts folding through one of these regions, folding proceeds through the other pathways with only minimal decrease in the folding rate (67). And only in the case of simultaneous disruption of all nucleation sites, a large decrease in the folding rate is observed. These observations are in the perfect consent with the DC model.

Taking together the diversity of the coiled coil class of proteins, complex nature of helix-stabilizing interactions and multiplicity of folding pathways within DC model, we suggest a reformulation of definite “THE triggering sequence” into an indefinite “A triggering site”, in accordance with the original DC microdomain concept. As shown below, GCN4p1 E22-R25 salt bridge is indeed not the solitary helix-stabilizing determinant in the GCN4p1 sequence.

### ***X-form***

Notwithstanding large amount of data pointing to the existence of a stable folding intermediates in the GCN4p1 folding pathway, for a long time no high-resolution experimental study has focused on their characterization. Recently, employing solution NMR methodology, we have identified and characterized the “x-form” – a novel stable conformation of GCN4p1, which exists in equilibrium with the coiled coil form (76). X-form is a semi-structured folding intermediate, populated at about 1% at ambient conditions, but considerably stabilized in the acidic pH. In the very first 1991 NMR structural study of GCN4p1 an additional set of resonances was observed at low protein concentration, pointing to the presence of a second conformational ensemble (93), which remained essentially unattended by the authors. All further high resolution studies were conducted either at the very high protein concentration (above 1 mM), or in the neutral pH range, where this novel conformational state (x-form) is only marginally populated. Concentration dependence of the x-form population together with the slow (millisecond) exchange regime between the coiled coil and x-form conformations, unequivocally renders the x-form as an intermediate at the monomer side of the folding barrier.

Experimental data on x-form is in good consent with the Diffusion-Collision theory. Within DC model folding proceeds via microdomains, which may be detected experimentally as structured intermediates transiently populated during the folding process. Populations of folding intermediates, as well as the final folded states, are strongly dependent on the kinetic rates of transitions between these conformations. X-form population increases in acidic pH and at low peptide concentration due to the perturbations of the kinetic rates of the transitions

linking x-form with dimeric structures. At the first stage of LZ folding helix propagation is energetically favorable (83), which leads to accumulation of x-form at the main transition barrier, unless x-form further associates to form the coiled coil dimer. When LZ peptide concentration is low, the rate of [successful] chain collisions decreases, reducing the rate of coiled coil formation and increasing the population of monomeric x-form intermediate.

Acidic pH appears to have additional double effect on the equilibrium distribution of the conformational states at the final stages of the LZ-GCN4 folding process. Firstly, protonation of the Glu side-chains eliminates the “electrostatic guidance” effect of attractive interchain ionic interactions (81), decreasing the apparent association rate from  $\sim 4$  (average from 6 studies) to  $0.08 [10^6 * M^{-1} * s^{-1}]$  (Table 4.1). Secondly, uncompensated positive charges in the final coiled coil structure destabilize the dimer, increasing the dissociation rate from 0.18 (average) to  $2.6 [s^{-1}]$  (Table 4.1). Combination of these effects leads to accumulation of the monomeric folding intermediate – x-form.

Structural information obtained from NMR data reveals a considerable amount of helical structure present in the x-form. Most importantly, these data indicate that LZ-GCN4 pre-collision intermediate bears two regions of increased helicity (76) – allocated in vicinity of intrahelical  $i, i+3$  salt bridges K8-E11 and E22-R25 (GCN4p1 numbering) (Figure 4.1), showing that E22-R25 “trigger site” is not the only helix-nucleating determinant in LZ-GCN4. Although it has been reported that in acidic pH the helix structure of the “triggering site” is nearly abolished (16), the side-chain  $pK_a$  studies show that none of the GCN4p1 glutamates is fully protonated at the pH 3.2 (94). This allows a fraction of salt-bridge stabilized helical microdomains to be maintained even at acidic conditions. Further high-resolution studies are required to elucidate the exact 3D structure of the x-form and explain different endurances of the N-terminal K8-E11 and C-terminal E22-R25 triggering sites.

### ***Importance of structured monomeric intermediates***

Although appearing marginally populated at high protein concentrations and neutral pH conditions, x-form could represent a considerable fraction of the LZ-GCN4 structural ensemble in the *in vivo* conditions of low peptide concentration. Therefore, this conformational state might be biologically relevant in providing a transient interface for recombination of LZ partners in the context of cellular signalling networks. Structured alpha-helical intermediates appear to be an essential part of the generalized LZ folding model. They provide a robust scaffold with native-like positioning of electrostatic and hydrophobic residues, which enables discrimination between individual interacting partners both on primary (sequence) and tertiary (chain register and orientation) structural levels.

### *Intermediates at the native side of the main folding barrier*

Even before first notions of monomeric helical intermediates have emerged, Hans Rudolf Bosshard group highlighted the biphasic nature of LZ folding kinetics (95). Soon after that Alfred Holtzer and colleagues employing equilibrium kinetic measurements by NMR have further challenged the apparent uniformity of the coiled coil ensemble at the native side of the transition barrier (88, 96-98). Initially these observations had been overwhelmed by experimental data from other groups, and only several years later more evidences of stable folding intermediates at the native side of the LZ folding barrier appeared (15, 89). The most intuitive picture on these folding transitions can be derived from a thermodynamic study by Anatoly Dragan and Peter Privalov (89). According to this study, unfolding of LZ-GCN4 can be modeled by at least three step mechanism, with first two transitions being concentration independent (unimolecular) and only third one – concentration-dependent dissociation of dimer strands. This indicates that at least three conformational ensembles exist at the native side of LZ folding barrier. The first transition, starting from a 100% coiled coil state occurs at the temperatures around 0°C for GCN4p1 and corresponds to the fraying of few N-terminal residues. The second transition occurs at much higher temperatures (starting at about ~20°C for GCN4p1) and shows perturbations at both termini as well as structure repacking in the central region of LZ. Further increase in temperature (above ~45°C for GCN4p1) eventually induces cooperative dissociation of coiled coil into monomers. Characteristics of the second structural transition, observed between 20 and 45°C before the dissociation of the dimer, coincide with the conformational exchange between two folded LZ states reported earlier by Alfred Holtzer and colleagues (88, 96). We dub the pre-dissociation dimeric intermediate state as a “relaxed coiled coil” conformation.

Ting Wang and coworkers, employing T-jump relaxation experiments on the cross-linked version of GCN4p1, reported two equilibrium processes with the time relaxation constants in the order of 100  $\mu$ s and 10  $\mu$ s (15). Authors attributed the slow (100  $\mu$ s) relaxation component to the bimolecular coiled coil folding reaction, and fast (10  $\mu$ s) component to the conformational exchange at the native side of the folding barrier. The observed timescale of the slow transition indeed fits perfectly to the timescale of GCN4p1 association. As shown by Liam Moran (67) 133  $\mu$ s ( $= 1/k_{on}$ ) folding timescale of the crosslinked GCN4p1 corresponds to ~5 ms for bimolecular association of 100  $\mu$ M non-crosslinked GCN4p1 ( $1/[100 \mu\text{M} * k_{on}]$ ) (Table 4.1). Therefore 100  $\mu$ s relaxation process reported by Wang fits to the millisecond timescale of GCN4p1 association reported by numerous research groups (compared at 100  $\mu$ M peptide concentration)(16, 74, 75, 80) (Table 4.1). However the fast (10  $\mu$ s) component assigned to unimolecular transitions at the native side of the folding barrier is in sharp contrast with the millisecond-to-second

timescale reported for GCN4p1 analogs (88, 96), as well as analogs of Jun-Fos heterodimer (78). This discrepancy is also highlighted by the differences in structural characteristics of reported states. The 10  $\mu$ s transition observed by Wang apparently involved a change in the secondary structure, while transitions observed by Andre d'Avignon and Jody Mason appeared to involve only repacking within the hydrophobic core while maintaining overall helical structure, in accordance with the “relaxed coiled coil” observed by Anatoly Dragan (89). Further equilibrium kinetic and thermodynamic studies involving dimeric (i.e. non-cross-linked) forms of GCN4p1 and its analogs are required to clarify the nature of this conformational state.

Of all interactions defining specificity and stability of canonical leucine zippers, only buried polar residues at the  $\alpha$ -positions of the interface have not been thoroughly investigated. It has been shown that these interactions do not affect the folding reaction rates prior to the main transition barrier (90), and therefore are considered to manifest themselves only at the dimeric side of the folding barrier. In this perspective it is tempting to speculate that the transition between the coiled coil and “relaxed coiled coil” states involves reorganization of the VdW packing and hydrogen bonding in vicinity of these buried polar sidechains. From this standpoint, buried non-hydrophobic sidechains would decrease the stability of the coiled coil state, reducing the height of the unfolding activation barrier and thereby modulating the lifetime of the LZ signalling event.

#### **= 4.4 = Diffusion-Collision-Desolvation (DCD) model**

Employing DC model LZ folding is defined by two transition barriers: minor helix nucleation barrier, which is traversed on the nanosecond timescale, followed by major dimerization barrier traversed on the millisecond timescale (Figure 4.2). Second barrier occurs upon the collision of monomer chains and reflects the probability of pre-formed structural elements to establish a productive transition state. Notwithstanding massive efforts invested in the kinetic and thermodynamic studies of LZ folding, some controversies remain in the view on the physical nature of the main folding barrier and transition state ensemble associated with this event. One of the prevailing concepts in LZ folding is that the transition state comprises poorly structured dimer, with transiently formed helices undergoing search of complementary nucleating segments via local VdW interactions (70, 74), and that inter-chain electrostatic interactions do not affect the early folding, appearing only at the native side of the transition barrier (80). In this representation, the nature of the main activation barrier in the folding direction is purely entropic.

On the other hand, it seems more logical to expect long-range Coulombic interactions to dominate over short range VdW and hydrophobic interactions in discrimination of

dimerization partners and stabilization of the transition state. Indeed, some facts support this point of view. First, the low buried nonpolar surface area (~10-30%) reported for the transition state ensemble in a few cases (74, 78) points to insufficiency of short range VdW interactions in establishing a productive transition state. This is supported by the notion of buried polar residues having negligible effect on LZ dimerization rates (90), an observation which is hard to envision if the transition state is stabilized by short-range interactions. As well several direct evidences point to the importance of electrostatic interactions within transition state ensemble. Hans Wendt in 1997 observed strong dependence of LZ folding rates on the ionic strength of the folding milieu, which suggested formation of an electrostatically stabilized transition intermediate during the rate-limiting step in the folding direction (68). In a more recent study, the folding rate of the Fos-JunW heterodimer decreased 6-fold upon introduction of additional charged residue by Q21R mutation, although the mutation increased local helical propensity and created a new interhelical salt bridge (78). Finally, group of Alfred Holtzer has shown that upon assembly of LZ-cJun dimer, nearly 65% of the free energy barrier in the folding direction is due to enthalpic contributions, seriously undermining the possibility of purely entropic nature of LZ transition state (77).

A similar controversy existed in discussions of importance of electrostatic contributions for the coiled coil dimer stability (23, 29, 38, 40-42), and has been recently resolved by Hans Rudolf Bosshard and Daniel Marti, showing that the net thermodynamic contribution of a salt bridge is balanced between favorable charge-charge interaction and unfavorable desolvation energy (43, 44). Same logic can be applied here to explain the electrostatic contribution to the main folding barrier. Namely, that favorable “electrostatic guidance” effect from complementary charges (81), while increasing the probability of productive transition ensemble formation, is often compensated by the slow desolvation of the involved charges (82). In other words, increasing the rate of folding due to electrostatic guidance is compensated by decrease of the same rate due to desolvation. In fact the impact of charge desolvation on the kinetics of LZ traversing the main folding barrier has been already discussed in a number of studies (77, 99, 100). In this perspective it appears reasonable to extend the Diffusion-Collision model of LZ folding to the Diffusion-Collision-Desolvation model, where LZ folding rate equation (2b) will be explicitly complemented by appropriate terms reflecting electrostatic (Coulombic) and desolvation contributions.

Besides resolving the aforementioned controversies on entropic vs. enthalpic nature of the transition barrier, these terms will rule out the argument used to question the validity of DC model in favor of the two-state mechanism. Tobin Sosnick and coworkers (67) argued that if the DC model was valid, 10-fold helicity drop in the A-peptide (66) compared to



GCN4p1 shall lead to a 100-fold drop in its folding rate, while in reality the peptide folded 20 times faster than the original sequence; and, similarly, engineered GCN4-E9G4 peptide according to DC was expected to fold 100- to 10000- fold slower, while in experiment it was 4-fold faster than original GCN4p1 (69). The main accent in interpretation of these results was made on the differences in helical propensities of the peptides, thus undermining the DC model. Meanwhile, along with the changes in helical propensities a serious perturbation in the Coulombic interactions was obviously introduced, enforcing unexpectedly high folding rates of uncharged GCN4p1 analogs compared to the wild-type sequence.

Finally, the DCD model resolves the argument of the kinetic (95, 101) versus thermodynamic (80) control of LZ specificity. As discussed above and in the previous section, the net contributions of charge-charge interactions to the coiled coil stability are relatively moderate (23, 41, 43, 44) and are generally overridden by free energies contributed through the residues at the hydrophobic core (Table 3.1). Under these conditions the control of LZ oligomerization specificity by interchain *g-e'* (*i, i+5*) ionic interactions (10, 22, 47) cannot be purely thermodynamic, since moderate thermodynamic stability conferred through ionic interactions would not be sufficient to discriminate specific oligomerization partners in the background of highly stable coiled coil core. Hence, the only time point where electrostatic interactions can effectively modulate oligomerization specificity are the early stages of LZ folding process, when VdW interactions and hydrophobic burial have not yet accreted the coiled coil structure.

Relating to the DCD model definition of LZ specificity, two general types of side-chains in the *g* and *e* positions can be considered: nonionic and ionic. Nonionic side-chains decrease specificity of LZ, allowing the peptide to indiscriminately interact both with charged and non-charged residues without significant perturbations of the activation barrier. Meanwhile ionic sidechains in *g/e* positions foster increased specificity of oligomerization, disfavoring partners with identically charged sidechains. Presence of repulsive interactions thus increases the activation barrier for folding, limiting the population of particular LZ oligomer, even though from thermodynamic standpoint oligomeric structure would be more favorable than separate monomer chains. Kinetic contribution to LZ specificity control is strongly corroborated by early observations that the LZ strand exchange is predominantly governed by the dissociation rate of the coiled coil dimer (95, 101), especially in the presence of its consensus DNA sequence (102, 103).

### ***Arguments against DCD***

One argument against the validity of the DCD model relies in the fact of additivity of the activation free energy  $\Delta\Delta G_{U \rightarrow \ddagger}^{\circ}$  (unfolded to transition state) perturbations obtained by

mutating cationic and anionic sidechains involved in interhelical salt bridges of GCN4p1 (80). This fact suggests that these interactions are not formed in the rate-limiting transition state. However, as shown by Hans Rudolf Bosshard (44), it is plausible that favorable effects from charge-charge interactions are compensated by unfavorable desolvation energies of the charged sidechains, leading to insignificant perturbation of the overall activation energy contribution.

Another argument can be found in the early study by the same Bosshard group, which shows a striking 75-fold difference between folding rates of almost identical leucine zippers Flu-LZ(12A) and Flu-LZ(16A), which differ only in positioning of a Leu>Ala mutation – being either at *d*-position (mutant 12A) or at *a*-position (mutant 16A). The difference cannot be easily accounted for neither by perturbations in the first (helix-priming) transition barrier, nor by electrostatically stabilized transition state of the second barrier. This indicates that the second folding barrier could be formulated both by short-range entropic and long-range enthalpic contributions. However this example has to be treated with caution due to the apparent complexity of the folding model employed – peptide under study had a tendency to trimerize at slightly higher concentrations, and exhibited a three-state behavior within the measurement timescale (95).

## **= 4.5 = Summary**

Summing up, the events governing the LZ folding can be described as following (Figure 4.2). The first, minor transition barrier, is associated with the helix nucleation event and is traversed on the nanosecond timescale (83-86). Here the helix-coil transition is favored by helical propensities (67, 70, 75, 104) and opposed by the losses in conformational entropy (104). Further helix propagation is energetically favorable (occurs down-hill), since each additional (*i*, *i*+4) hydrogen bond entropically restrains only one residue, while preceding nucleation event requires simultaneous fixation of three residues (83). Depending on the particular sequence, preformed helical intermediates can be additionally stabilized by intramolecular (*i*, *i*+3) electrostatic interactions (16, 78, 80), as exemplified by E22(*g*)-R25(*c*) “triggering site” observed in GCN4p1. Overall this initial step leads to formation of partially structured monomeric intermediates, which represent an association-competent state capable of recognizing specific oligomerization partners.

Subsequent period is limited by diffusion and is traversed on the microsecond timescale (Table 4.1) (66, 75, 87). It is followed by the second, major energetic barrier, which occurs upon the collision of monomer chains and reflects the ability of pre-formed structural intermediates in establishing a productive transition state. In the forward (folding) direction, this barrier is traversed on the millisecond timescale (Table 4.1) (16, 67, 72, 74, 75, 80) and is

predominantly modulated by establishment of interhelical ionic interactions. Enthalpic nature of this barrier provides a certain flexibility in definition of its activation energy, stemming from interplay between the attractive and repulsive “electrostatic guidance” (81) effects and energetic penalty associated with charge desolvation (44, 82). This flexibility provides one of the major tools for control over specificity of LZ oligomerization. Establishment of the productive transition state is followed by a rapid down-hill “zipping up” of the structure, associated with stabilization of coiled coil structure by short-range VdW interactions (71, 90), burial of the hydrophobic core (66, 67, 71) and formation of remaining ionic interactions (74, 80, 87).

The traverse of the main energetic barrier in the reverse (unfolding) direction occurs on the scale of seconds and even minutes (Table 4.1). This transition is characterized by cooperative dissociation and unfolding of monomer chains, driven by favorable increase in conformational entropy (74, 77, 87, 89). The activation energy of this transition is dependent both on the properties of the hydrophobic core (66, 67, 71, 90) and interchain electrostatic interactions (43, 74, 80, 87).

Reflected in this summary, our knowledge regarding folding of dimeric coiled coils, and leucine zippers in particular, predominantly relies on the studies of LZ-GCN4 (GCN4-p1) and its analogs. Although providing a common reference frame for different research groups, in some aspects this tailored approach turned out counter-productive. This is highlighted by delayed recognition of enthalpic contributions to the main folding barrier, which became apparent only after studies were extended to include Jun and Fos LZ motifs. At the moment, stability and specificity rules described in the previous section, together with the knowledge of LZ folding landscape allow approximation of experimental findings to other LZ motifs and short coiled coil sequences. For example, compared to LZ-GCN4, coiled coil dimer of human LZ-cJun is considerably less stable, with over five orders magnitude difference in their equilibrium unfolding (dissociation) constants ( $K_d$  446 $\mu$ M vs. ~2-8 nM) (77). This instability can be attributed to the diminished amount of favorable ***a-a'*** interactions (50% vs. 85% in LZ-GCN4) and absence of favorable electrostatic ***g-e'*** couplings (0 vs. 3 interhelical ***g-e'*** salt bridges in LZ-GCN4) (77). In addition LZ-cJun shows a considerable decrease in the folding rate. Latter effect is likely dictated by the reduced amount of stabilizing intra-helical salt bridges (1 vs. 3), absence of attractive inter-helical ***g-e'*** ionic interactions (0 vs. 3) and presence of one repulsive inter-helical ***g-e'*** interaction (Lys7 – Lys12). Within the proposed Diffusion-Collision-Desolvation model these differences manifest themselves respectively by decreasing the stability of helical pre-collision intermediates, eliminating favorable

electrostatic guidance of monomer chains and decreasing the amount of successful collision events.

## **= 4.6 = Conclusion**

For a long time considered a simple two-state transition, folding of leucine zippers appears a rather complex process, very sensitive to alterations in the peptide composition and experimental conditions. Accordingly, specific kinetic and thermodynamic characteristics of LZ folding landscape will be dependant on the exact peptide sequence and the folding environment.

Importantly, as demonstrated by the diversity of amino acids in the *g*, *e* and *a* positions of the leucine zipper interfaces (7, 10), Leucine Zipper motifs have emerged under specificity-driven, rather than stability-driven evolutionary pressures. Combinatorial multiplicity of LZ interfaces together with kinetic flexibility of their folding landscape, provide the fundamental basis for their versatility in establishment of protein signalling cascades and networks.

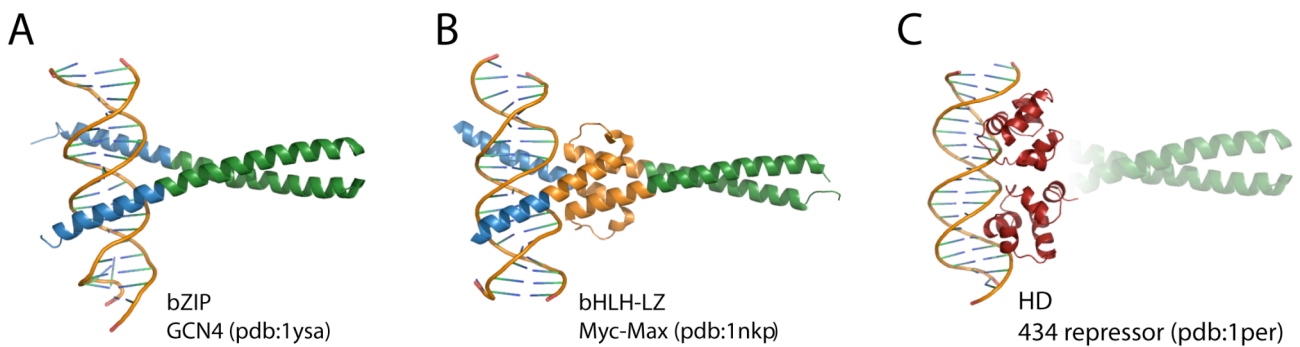
## **= 5 = Functional diversity**

In the previous sections we have reviewed the knowledge on LZ structure, specificity and folding properties, revealing the molecular mechanisms of signal transduction by these motifs. Leucine Zipper signalling is based on a shared set of specificity determinants exposed within a common coiled coil scaffold, thus enabling sophisticated combinatorial protein signalling networks. As will be shown in this section LZ networks cover a diverse set of regulatory pathways and the universal LZ “interaction code” indeed appears to provide a common framework for interconnecting various signalling cascades. Currently no appropriate analytical tools are available for precise quantitative analysis and prediction of LZ motif interactions. Nonetheless, even qualitative “LZ code” analysis using empirical rules discussed above, already permits predictions to be made from the primary sequences of LZ motifs. Beyond widely known bZIP, bHLH-LZ and HD-ZIP protein networks, several other examples of LZ-mediated signalling pathways are presented, along with evidences of couplings between different pathways enabled via LZ motifs.

## = 5.1 = Transcription factors – bZIP, bHLH-LZ, HD-ZIP

### *Origins*

The most widely known families of LZ-containing proteins are basic region leucine zipper (bZIP, Figure 5-1A) (12), basic region helix-loop-helix leucine zipper (bHLH-LZ, Figure 5-1B) (105) and homeodomain (aka helix-turn-helix) leucine zipper factors (HD-ZIP, apparently no native HD-ZIP structure is available to date, so a tentative structure is shown in Figure 5-1C) (106). Underscoring universality of LZ interaction, several distinct LZ transcription factor networks have emerged in different lineages at different periods of evolution. For example, bZIP and bHLH-LZ families emerged around 1.6 billion years ago and are common for plant, fungi and animal kingdoms, meanwhile HD-ZIPs emerged independently in the plant lineage around 0.7 BYA (25).



**Figure 5.1.** Structure of leucine zipper transcription factors. (A) bZIP GCN4 homodimer (pdb:1ysa). (B) bHLH-LZ Myc-Max heterodimer (pdb:1nkp). (C) Hypothetical structure of HD-ZIP factor. Homeodomain from 434 repressor protein (pdb:1per) shown along with tentative positioning of LZ motif. Leucine zipper motifs marked green, basic region – blue, helix-loop-helix (B) – orange, homeodomain (C) – dark red.

### *Structure*

In all three LZ-TF families the leucine zipper is positioned C-terminal to the DNA-binding motif. In the case of bZIP and bHLH-LZ specific DNA-binding determinant is represented by the basic region of the molecule, and in HD-ZIP this function is mainly conferred by the third helix of the helix-turn-helix motif. While bZIP and HD-ZIP possess “two-domain” [basic region / homeodomain] + [leucine zipper] architecture, bHLH-LZ (and bHLH in general) possess “three-domain”, sometimes even “four-domain” architecture: [basic region] + [helix-loop-helix] + [leucine zipper / orange / ...] (25). This permits a more versatile network of interactions to be formed in the latter case, which is reflected in the size of bHLH family (25). Another adaptation for increased number of regulatory pathways is

found in LZ motifs of plant bZIPs – on average those are  $\geq 8$  heptads long, compared to 4-6 heptads of human bZIPs, thus considerably extending their combinatorial specificity (10).

### ***Functions***

Leucine zipper transcription factors evolved as key regulators in a wide variety of processes. Today they are truly widespread among eukarya with only human genome encoding 51 proteins with unique bZIP motifs (107) and at least 31 proteins with unique bHLH-ZIP motifs (108). In Arabidopsis similar bZIP / bHLH-LZ array is complemented with 47 unique HD-ZIP proteins (25).

bZIP is the second-largest family of dimerizing transcription factors in humans (25). They control expression of genes involved in development, environmental stresses, metabolism, circadian clock and neuronal activity (26), with a number of factors being widely renowned oncogenes, such as AP-1 (109). Similarly, in plants bZIPs mediate diverse developmental and homeostatic processes, as well acting in various environmental stress signalling pathways (110).

bHLH-LZ represent a subset of bHLH proteins – the largest family of dimerizing transcription factors across eukarya (25). The key targets of bHLH-LZ regulation are developmental processes, differentiation and, most importantly, cell cycle (111, 112). Similarly to bZIP family, dysfunctions of bHLH-LZs are strongly associated with tumorigenesis, with the most prominent example being highly oncogenic transcription factor c-Myc (113), which regulates up to 15% of all genes in an organism (114).

HD-ZIP factors are unique to the plant kingdom, and, as opposed to the archetypical homeodomain proteins, do not play homeotic roles, exhibiting highly specialized functionality (115). These factors are split into four classes based on structural characteristics, and govern a spectrum of processes including responses to environmental conditions, hormone action, meristem regulation and organ development (115).

### ***“LZ silencing” and pharmacological implications***

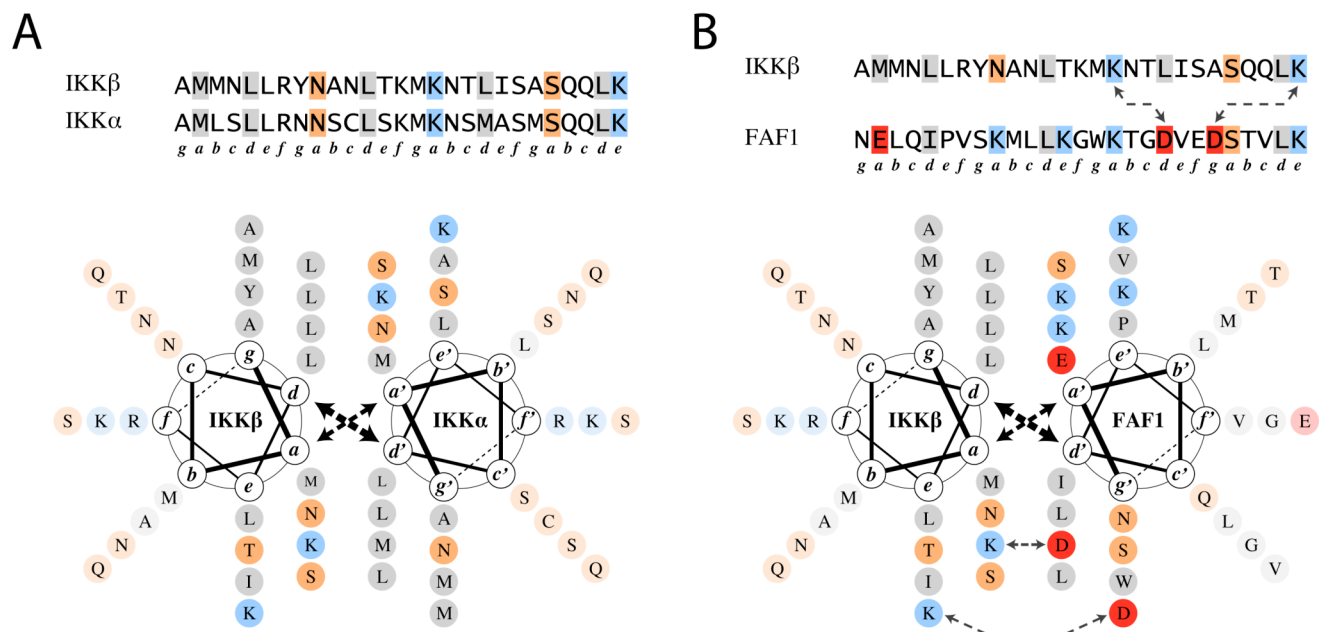
As proposed by Grigoris Amoutzias and others, evolution of organisms starting from eukarya to a large extent depended on the expansion of TF networks, those including bHLH, bZIP, NR and MADS (25, 108). Emerging complexity of these networks in turn stimulated formation of additional regulatory machinery, in case of LZ proteins involving cofactor interactions (116, 117), compartmentalization (118), proteolysis (119, 120), phosphorylation (121, 122), acetylation (123), glycosylation (124) and etc (reviewed for plant bZIPs in (125)).

One of the most important regulatory mechanisms within the networks of dimerizing TF is suppression of signal transduction in presence of complementary dimerization partners

lacking DNA-binding motifs. This modulation, targeted towards the key hubs of the TF network, provides a mechanism for robust silencing of extensive network segments. One of the most renowned examples in this perspective are Id proteins (“inhibitor of DNA binding”) of bHLH family, which employ dominant negative helix-loop-helix interactions to regulate cell cycle, tumorigenesis and cellular senescence (126, 127). As for the Leucine Zipper transcription factors, a similar example has been recently reported in regulation of HD-ZIPs – a ZPR family represents short LZ-peptides which specifically heterodimerize with HD-ZIPIII transcription factors impairing their DNA-binding abilities (128). Along these lines, a number of drug-development studies are under way focusing on targeted silencing of LZ-mediated signal transduction by small molecules (129-131) as well as complementary dominant-negative LZ motifs (132-134).

## = 5.2 = Immune response signalling – NF-kappaB pathway

Leucine zipper interactions play an important role in the cytoplasmic part of antiapoptotic NFκB signalling pathway, which is conserved from flies to mammals and provides one of the key routes for the immune and stress responses (reviewed in (135)). Many of the extracellular signals that lead to the activation of NFκB converge on a ~900 kDa IκB kinase complex (IKK) consisting of three major subunits – catalytic IKKα and IKKβ (52% identical in primary structures), and regulatory IKKγ (also known as NEMO) (136-138) or reviewed in (139) and (140).

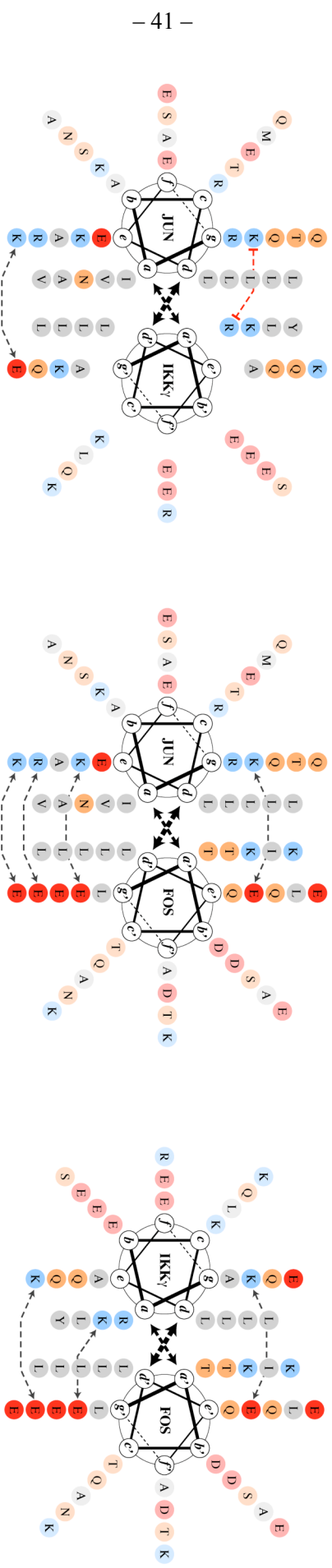
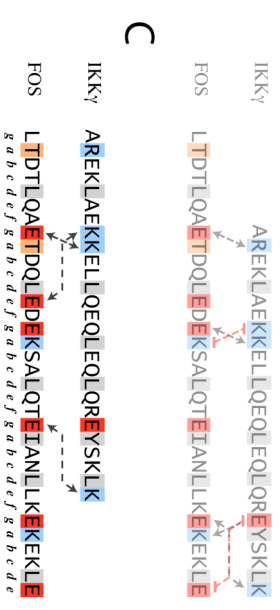
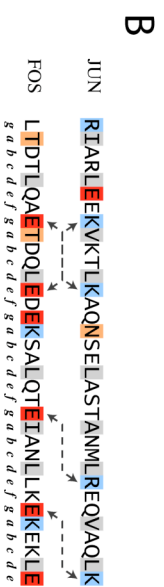


**Figure 5.2.** Predicted interactions of LZ motifs from factors IKKα/IKKβ (A) and IKKβ/FAF1 (B).

LZ motifs of both catalytic IKK $\alpha$ /IKK $\beta$  subunits exhibit identical polar residues in all four **a**-positions of the dimer interface (Met, Asn, Lys, Ser - **Figure 5.2-A**), which provides a very specific oligomerization determinant. Formation of IKK $\alpha$ /IKK $\beta$  heterodimer is indispensable for the activation of the whole complex and its kinase activity (*137, 141*).

The regulatory subunit IKK $\gamma$  interacts with the catalytic IKK $\alpha$ /IKK $\beta$  heterodimer via its N-terminal coiled coil motif, while the C-terminal LZ motif of IKK $\gamma$  plays a more intriguing role. As has been shown by Marshall Horwitz group – LZ of IKK $\gamma$  enables direct interaction of the IKK complex with the components of the AP-1 complex (c-Jun and c-Fos) (*121*). Thermodynamic analysis of heterodimeric LZ interactions in the IKK $\gamma$ •Jun versus IKK $\gamma$ •Fos complexes employing the simplified set of LZ determinants (**a-a'**, **d-d'** and **g-e'**) shows that both complexes are comparably stable and specific. However analysis of the corresponding wheel diagrams shows that kinetically IKK $\gamma$ •Jun complex is less favorable, forming one attractive and one repulsive interchain ionic interaction, opposed to 3 attractive ionic interactions in IKK $\gamma$ •Fos complex (**Figure 5.3**). Therefore, a more precise contextual analysis of sidechain interactions within these complexes is required to make accurate predictions of AP-1 (Jun•Fos) equilibrium redistribution in the presence of IKK $\gamma$ . In any case, the discovery of interactions between IKK $\gamma$  and AP-1 sheds some light on the debated issue of molecular mechanisms of coupling between the AP-1 and NF-kappaB pathways and delicate balance between cell death and cell survival decisions (*121*). It also provides an illustrative example of LZ-mediated coupling between cytoplasmic and nuclear protein signalling pathways.





**Figure 5.3.** Predicted interactions between LZ motifs from factors Jun•IKK $\gamma$  (A), Jun•Fos (B) and IKK $\gamma$ •Fos (C). Linear representations show two alternative alignments of the 5-heptad Jun and Fos zippers with the 4-heptad IKK $\gamma$  LZ motif. Thermodynamically unfavorable alignments are dimmed. Wheel representations illustrate energetically most favorable alignments of the LZ motifs.



the protein shall be highly stable as a homodimer – containing 6 favorable intermolecular *g-e'* salt bridges, all-Leu in *d*-positions and predominantly beta-branched residues in *a*-positions (Figure 5.3). Presence of two polar buried asparagines increases the homodimerization specificity even further. According to experimental data LZ-mediated GILZ homodimerization is indeed essential for its function as inhibitor of NF- $\kappa$ B transcriptional activity (145).

### **= 5.3 = More kinases – PKG, ZIPK, DAPK**

Similar to the NF-kappaB example, there are many other seemingly distinct kinase signalling pathways, possibly coupled through the common structural nature of LZ interfaces.

Regulation of vascular smooth muscle tone is mediated through the balance of myosin phosphorylation and dephosphorylation rates. In the case of nitric oxide dependant vasodilation, this involves a multiplex LZ-mediated equilibrium between homodimeric form of PKG-I $\alpha$  (cGMP-dependent protein kinase I $\alpha$ ) and its heterotetrameric assembly with the MYPT1 (myosin-binding subunit of the myosin phosphatase) (146, 147).

Similarly, another enzyme – Zipper-interacting protein kinase (ZIPK) modulates the phosphorylation state of myosin light chains in the case of smooth muscle contraction in response to Ca<sup>2+</sup> (148). At the same time, ZIPK is known as the death-associated protein kinase 3 (DAPK3), enabling cell death through apoptosis (149). ZIPK leucine zipper motif was shown indispensable for its enzymatic activity (149), cellular localization and proapoptotic effects (150). Moreover, ZIPK LZ motif facilitates its heterodimerization with the transcription factor 4 (ATF4) from the bZIP ATF/CREB family (26) through their LZ domains (149), providing direct coupling between nuclear and cytoplasmic signalling networks.

### **= 5.4 = Ion channels – AKAP**

Shortly after original discovery of LZ in transcription factors, the presence of this motif was revealed in the family of voltage-gated potassium channels with suggested involvement in subunit interactions, mediating voltage-dependant opening and closing of the channel (13). In skeletal and cardiac myocytes interaction of the LZ motifs of L-type Ca<sup>2+</sup> channels (CaV1.1 and 1.2) and A kinase-anchoring protein (AKAP15) provides an efficient mechanism for anchoring of cAMP-dependent protein kinase (PKA) to Ca<sup>2+</sup> channels, ensuring rapid and efficient phosphorylation of Ca<sup>2+</sup> channels in response to local signals such as cAMP and depolarization (151, 152). This mechanism is very similar to the hippocampal pyramidal cells, where rapid modulation of neuronal excitability through Na<sup>+</sup>

channels occurs by local protein phosphorylation of the channel by the protein kinase A specifically recruited via its LZ motif by the A kinase-anchoring protein AKAP15(*153*).

### **= 5.5 = Transport vesicles – SNARE**

Many biomolecules are transferred among different cellular compartments by transport vesicles, which recognize and merge with the target compartments in a highly specific manner in addition overcoming a very high activation energy barrier during membrane fusion (nicely summarized in (*134*)). In all eukaryotic cells this task is accomplished by a family of proteins called SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) (*154*). All SNAREs bear a conserved cytosolic coiled-coil/leucine zipper motif of 60-70 residues, which assembles into a parallel four-helix bundle – SNAREpin. Structural classification of SNAREs is based on the type of sidechain they contribute to the “zero ionic layer” – a cluster of buried ionic interactions which apparently define specificity of the interface (*155, 156*).

Essentially SNARE is a mixed “leucine zipper” (signal) / “coiled coil” (structure) motif, serving two goals – four-helix-bundle provides a highly stable interface providing energy for membrane fusion, while “ionic zero layer” creates a unique specificity determinant, a functional analog of the buried polar residues defining specificity within bZIP signalling networks (see section 3.3). As one could anticipate, in SNARE tetramer the specificity determinant is located in the *d*-position of interface, adopting the same “perpendicular” geometry in tetrameric LZ as the *a*-positioned sidechains do in the dimeric leucine zippers (see section 2.2). Thus SNARE represents an intermediate state between the structural coiled coil and signalling LZ motifs – extreme stability of the four-helix-bundle allows “signal transmission” only in the forward direction, limiting thermodynamic control of the signalling event and requiring external factors for the SNARE tetramer disassembly (*134*).

### **= 5.6 = Viral envelopes and capsides**

Another discovery made shortly after the original William Landschultz publication on leucine zippers – the notion of paramyxovirus fusion glycoproteins facilitating dimerization and tetramerization via LZ motifs (*14*). Nowadays a wealth of experimental evidence exists signifying importance of LZ in the formation of viral capsides. For example HIV-1 envelope glycoprotein gp41 bears two LZ motifs – in its N- and C- terminal regions. Experimental results suggest that the corresponding N-leucine zipper, along with N-terminal fusion and viral membrane-adjacent regions of HIV-1 gp41 may promote key membrane perturbations underlying the merging of the viral envelope with the cell surface (*157*). Mason-Pfizer monkey virus (M-PMV) Gag protein contains a domain p12 that is unique to this virus

(simian retrovirus-3) and its close relatives. This domain incorporates a leucine zipper-like region that facilitates Gag-Gag protein oligomerization performing scaffold-like function within the viral envelope (*158*). Encapsidation of the herpes simplex virus type 1 (HSV-1) relies on the formation of 12-subunit ring structure mediated by LZ motifs of UL6 protein (*159*).

Viral LZ motifs are not directly involved in processing of target-recognition signals, frequently conducting pure mechanistic functions like membrane fusion. Nevertheless, as will be discussed below, nucleocapsid LZ motifs may serve as determinants specifically recognized by the innate antiviral immunity systems of higher eukaryotes.

### **= 5.7 = Innate antiviral defense – interferon induced Mx proteins**

Mx proteins are induced by alpha/beta interferons, forming one of the key components of the innate immune response against RNA viruses (*160*). Featuring highly conserved N-terminal GTPase domain, they are classified together with the dynamin-like large guanosine triphosphatases (GTPases), known to be involved in intracellular vesicle trafficking and organelle homeostasis. Beyond GTPase domain the sequence similarity fades away and at the C-terminus Mx proteins carry a unique LZ doublet (*161*), which replaces the PH-GED-PRD (Pleckstrin Homology + GTPase Effector + Proline-Rich Domain) triad characteristic for dynamins (*160*). This extended LZ motif empowers Mx proteins with antiviral activity against a wide spectrum of viruses, including members of bunyaviridae, orthomyxoviridae, paramyxoviridae, rhabdoviridae, togaviridae, picornaviridae, and Hepatitis B virus (*160*).

The importance of the Mx system was effectively demonstrated by Peter Stäheli and Otto Haller groups on the mouse models: disruption of a single Mx1 gene causes complete loss of innate immunity against mouse-adapted influenza, leading to an overwhelming infection and rapid death (*162, 163*). While enhanced Mx1 production efficiently protects transgenic mice against pandemic human 1918 influenza virus and highly lethal H5N1 strains (*163*).

It is proposed that Mx GTPases detect viral infection by sensing nucleocapsid-like structures, trapping viral components and making them unavailable for the generation of new viral particles (*164*). Being antivirally active as monomers, Mx proteins assemble into high-molecular-weight oligomers in the solution, possibly yielding a stable intracellular pool from which individual monomers are recruited in the presence of viral particles (*165, 166*). The detailed mechanism of Mx oligomerization is poorly understood, but is likely mediated by intermolecular domain swapping involving C-terminal LZ and coiled coil motif within central CID domain (*165, 167*). Dependence of Mx oligomerization on intramolecular backfolding of C-terminal LZ onto CID coiled coil motif remains controversial (*167*).

In terms of interaction determinants, the C-terminal leucine zipper motifs appear to be crucial for Mx antiviral activity (165, 168). As shown by mutagenesis studies, Mx leucine zipper represents a multipurpose recognition motif, shaped for identification of a diverse array of viral species. For example a point E645R mutation (at the *f*-position of the coiled coil) reshapes the interaction surface of human MxA, impairing recognition of the La Crosse (169) and Crimean-Congo hemorrhagic fever (170) viruses from the Bunyaviridae family. At the same time this E645R mutation does not affect Mx sensitivity to the Influenza A and Thogoto viruses from Orthomyxoviridae family (168, 169). Notwithstanding recent experimental advances, the molecular mechanism of Mx antiviral activity is still poorly understood. Particularly, the exact structural determinants responsible for the recognition of virus components have yet to be reported. Nonetheless, the dependence of Mx antiviral activity on integrity of its LZ motifs, together with high abundance of coiled coil structures in viral capsids and nucleoproteins, suggests that specific LZ interactions might be the key behind innate antiviral immunity response mediated by Mx proteins.

## = 6 = LZ in protein engineering

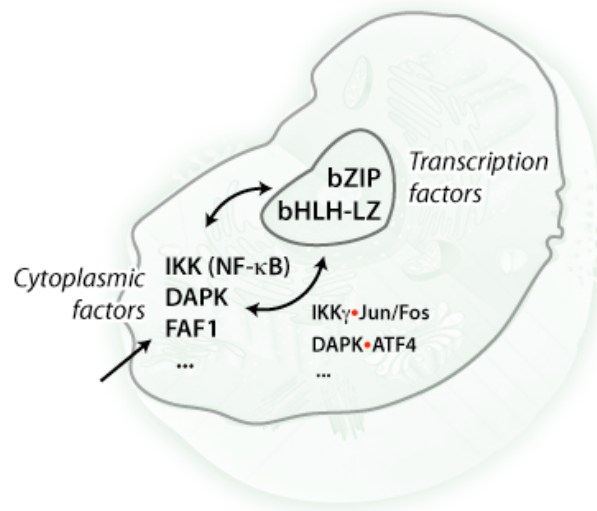
*under construction ...*

## = 7 = Conclusions and Outlook

The distinction between structural Coiled Coil motifs and signal-transducing Leucine Zipper motifs, like in the case of SNAREs, might sometimes remain elusive. Nevertheless, specific interactions enabled by short coiled coils with the hydrophobic core comprised predominantly of Leu sidechains, makes a notable distinction for the Leucine Zipper protein motif. We propose that on the cellular level uniformity of LZ interfaces enables one of the key components of the “protein interactome”, which facilitates coupling of many functionally distinct signalling pathways into one signalling network (Figure 5.5).

Further developments in understanding of the “LZ interaction code” require more thorough sampling of natural proteins for LZ motifs with new types of specificity determinants; as well as systemic engineering approaches for acquisition of precise physico-chemical properties enabling stability and specificity variation within LZ scaffold. As exemplified by the *a–d* ionic bridges in the Myc/Max network (36) and conserved *f*-position prolines found in some plant bZIP families (64), specificity rules derived from the studies of human bZIPs provide only a fraction of the picture.

Another important direction of research is development of molecular models and prediction tools for analysis of transient LZ interfaces. Currently available contextual data is insufficient for accurate prediction of stability and specificity of particular LZ interactions, even at the level of most commonly reported interaction determinants (*a-a'*, *d-d'*, *g-e'*). Extended vocabulary of specificity determinants, together with accurate prediction algorithms and intuitive visualization models shall facilitate a deeper understanding of functional implications generated by LZ signalling networks.



**Figure 5.5.** Schematic illustration of possible coupling between nuclear and cytoplasmic signalling pathways, enabled by LZ-proteins.

## = 8 = References

1. Walshaw, J., and Woolfson, D. N. (2001) Socket: a program for identifying and analysing coiled-coil motifs within protein structures, *J Mol Biol* 307, 1427-1450.
2. Lupas, A. N., and Gruber, M. (2005) The structure of alpha-helical coiled coils, *Advances in protein chemistry* 70, 37-78.
3. Lumb, K. J., Carr, C. M., and Kim, P. S. (1994) Subdomain folding of the coiled coil leucine zipper from the bZIP transcriptional activator GCN4, *Biochemistry* 33, 7361-7367.
4. Perera, R., Owen, K. E., Tellinghuisen, T. L., Gorbalenya, A. E., and Kuhn, R. J. (2001) Alphavirus nucleocapsid protein contains a putative coiled coil alpha-helix important for core assembly, *J Virol* 75, 1-10.
5. Burkhard, P., Meier, M., and Lustig, A. (2000) Design of a minimal protein oligomerization domain by a structural approach, *Protein Sci* 9, 2294-2301.
6. Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Engel, J., Aepli, U., and Steinmetz, M. O. (1998) An autonomous folding unit mediates the assembly of two-stranded coiled coils, *Proc Natl Acad Sci U S A* 95, 13419-13424.
7. Acharya, A., Rishi, V., and Vinson, C. (2006) Stability of 100 homo and heterotypic coiled-coil a-a' pairs for ten amino acids (A, L, I, V, N, K, S, T, E, and R), *Biochemistry* 45, 11324-11332.
8. Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001) Coiled coils: a highly versatile protein folding motif, *Trends in cell biology* 11, 82-88.
9. Woolfson, D. N. (2005) The design of coiled-coil structures and assemblies, *Advances in protein chemistry* 70, 79-112.
10. Vinson, C., Acharya, A., and Taparowsky, E. J. (2006) Deciphering B-ZIP transcription factor interactions in vitro and in vivo, *Biochimica Et Biophysica Acta- Gene Structure and Expression* 1759, 4-12.
11. Grigoryan, G., and Keating, A. E. (2008) Structural specificity in coiled-coil interactions, *Curr Opin Struct Biol*.
12. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins, *Science* 240, 1759-1764.
13. McCormack, K., CAMPANELLI, J., RAMASWAMI, M., MATHEW, M., TANOUYE, M., IVERSON, L., and RUDY, B. (1989) Leucine-zipper motif update, *Nature* 340, 103-104.
14. Buckland, R., and Wild, F. (1989) Leucine zipper motif extends, *Nature* 338, 547.
15. Wang, T., Lau, W. L., DeGrado, W. F., and Gai, F. (2005) T-jump infrared study of the folding mechanism of coiled-coil GCN4-p1, *Biophys J* 89, 4180-4187.
16. Steinmetz, M. O., Jelesarov, I., Matousek, W. M., Honnappa, S., Jahnke, W., Missimer, J. H., Frank, S., Alexandrescu, A. T., and Kammerer, R. A. (2007) Molecular basis of coiled-coil formation, *Proc Natl Acad Sci U S A*.
17. Rieker, J. D., and Hu, J. C. (2000) Molecular applications of fusions to leucine zippers, *Methods Enzymol* 328, 282-296.
18. Liu, J., Zheng, Q., Deng, Y., Cheng, C. S., Kallenbach, N. R., and Lu, M. (2006) A seven-helix coiled coil, *Proc Natl Acad Sci U S A*.
19. McLachlan, A. D., and Stewart, M. (1975) Tropomyosin coiled-coil interactions: evidence for an unstaggered structure, *J Mol Biol* 98, 293-304.
20. Crick, F. H. C. (1953) The Packing of Alpha-Helices - Simple Coiled-Coils, *Acta Crystallogr* 6, 689-697.
21. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil, *Science* 254, 539-544.



22. Vinson, C. R., Hai, T., and Boyd, S. M. (1993) Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design, *Genes Dev* 7, 1047-1058.
23. Krylov, D., Mikhailenko, I., and Vinson, C. (1994) A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions, *Embo J* 13, 2849-2861.
24. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants, *Science* 262, 1401-1407.
25. Amoutzias, G. D., Robertson, D. L., Van de Peer, Y., and Oliver, S. G. (2008) Choose your partners: dimerization in eukaryotic transcription factors, *Trends Biochem Sci* 33, 220-229.
26. Deppmann, C. D., Alvania, R. S., and Taparowsky, E. J. (2006) Cross-species annotation of basic leucine zipper factor interactions: Insight into the evolution of closed interaction networks, *Mol Biol Evol* 23, 1480-1492.
27. Serrano, L., Horovitz, A., Avron, B., Bycroft, M., and Fersht, A. R. (1990) Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles, *Biochemistry* 29, 9343-9352.
28. Moitra, J., Szilak, 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.
29. 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, *J Mol Biol* 279, 959-972.
30. Acharya, A., Ruvinov, S. B., Gal, J., Moll, J. R., and Vinson, C. (2002) A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: amino acids I, V, L, N, A, and K, *Biochemistry* 41, 14122-14131.
31. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1998) Orientation, positional, additivity, and oligomerization-state effects of interhelical ion pairs in alpha-helical coiled-coils, *J Mol Biol* 283, 993-1012.
32. Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999) The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins, *Protein Sci* 8, 2312-2329.
33. Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000) Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d", *J Mol Biol* 300, 377-402.
34. Cohen, C., and Parry, D. A. (1990) Alpha-helical coiled coils and bundles: how to design an alpha-helical protein, *Proteins* 7, 1-15.
35. Lavigne, P., Crump, M. P., Gagne, S. M., Hodges, R. S., Kay, C. M., and Sykes, B. D. (1998) Insights into the mechanism of heterodimerization from the 1H-NMR solution structure of the c-Myc-Max heterodimeric leucine zipper, *J Mol Biol* 281, 165-181.
36. Montagne, M., Naud, J. F., and Lavigne, P. (2008) Elucidation of the structural determinants responsible for the specific formation of heterodimeric Mxd1/Max b-HLH-LZ and its binding to E-box sequences, *J Mol Biol* 376, 141-152.
37. Alber, T. (1992) Structure of the leucine zipper, *Curr Opin Genet Dev* 2, 205-210.
38. Marti, D. N., Jelesarov, I., and Bosshard, H. R. (2000) Interhelical ion pairing in coiled coils: solution structure of a heterodimeric leucine zipper and determination of pKa values of Glu side chains, *Biochemistry* 39, 12804-12818.
39. Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins, *Faseb J* 10, 75-83.

40. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994) The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability, *Protein Eng* 7, 1365-1372.
41. Lumb, K. J., and Kim, P. S. (1995) Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper, *Science* 268, 436-439.
42. Phelan, P., Gorfe, A. A., Jelesarov, I., Marti, D. N., Warwicker, J., and Bosshard, H. R. (2002) Salt bridges destabilize a leucine zipper designed for maximized ion pairing between helices, *Biochemistry* 41, 2998-3008.
43. Marti, D. N., and Bosshard, H. R. (2003) Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges, *J Mol Biol* 330, 621-637.
44. Bosshard, H. R., Marti, D. N., and Jelesarov, I. (2004) Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings, *J Mol Recognit* 17, 1-16.
45. Moll, J. R., Olive, M., and Vinson, C. (2000) Attractive interhelical electrostatic interactions in the proline- and acidic-rich region (PAR) leucine zipper subfamily preclude heterodimerization with other basic leucine zipper subfamilies, *J Biol Chem* 275, 34826-34832.
46. Vinson, C., Myakishev, M., Acharya, A., Mir, A. A., Moll, J. R., and Bonovich, M. (2002) Classification of human B-ZIP proteins based on dimerization properties, *Mol Cell Biol* 22, 6321-6335.
47. Deppmann, C. D., Acharya, A., Rishi, V., Wobbles, B., Smeekens, S., Taparowsky, E. J., and Vinson, C. (2004) Dimerization specificity of all 67 B-ZIP motifs in *Arabidopsis thaliana*: a comparison to *Homo sapiens* B-ZIP motifs, *Nucleic Acids Res* 32, 3435-3445.
48. Alberti, S., Oehler, S., von Wilcken-Bergmann, B., and Muller-Hill, B. (1993) Genetic analysis of the leucine heptad repeats of Lac repressor: evidence for a 4-helical bundle, *Embo J* 12, 3227-3236.
49. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) Electrostatic interactions control the parallel and antiparallel orientation of alpha-helical chains in two-stranded alpha-helical coiled-coils, *Biochemistry* 33, 3862-3871.
50. Yadav, M. K., Leman, L. J., Price, D. J., Brooks, C. L., 3rd, Stout, C. D., and Ghadiri, M. R. (2006) Coiled coils at the edge of configurational heterogeneity. Structural analyses of parallel and antiparallel homotetrameric coiled coils reveal configurational sensitivity to a single solvent-exposed amino acid substitution, *Biochemistry* 45, 4463-4473.
51. Liu, J., Yong, W., Deng, Y., Kallenbach, N. R., and Lu, M. (2004) Atomic structure of a tryptophan-zipper pentamer, *Proc Natl Acad Sci U S A* 101, 16156-16161.
52. Zeng, X., Zhu, H., Lashuel, H. A., and Hu, J. C. (1997) Oligomerization properties of GCN4 leucine zipper e and g position mutants, *Protein Sci* 6, 2218-2226.
53. Zeng, X., Herndon, A. M., and Hu, J. C. (1997) Buried asparagines determine the dimerization specificities of leucine zipper mutants, *Proc Natl Acad Sci U S A* 94, 3673-3678.
54. Holton, J., and Alber, T. (2004) Automated protein crystal structure determination using ELVES, *Proc Natl Acad Sci U S A* 101, 1537-1542.
55. Yoon, M. K., Kim, H. M., Choi, G., Lee, J. O., and Choi, B. S. (2007) Structural basis for the conformational integrity of the *Arabidopsis thaliana* HY5 leucine zipper homodimer, *J Biol Chem* 282, 12989-13002.
56. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) Buried polar residues and structural specificity in the GCN4 leucine zipper, *Nat Struct Biol* 3, 1011-1018.
57. Oakley, M. G., and Hollenbeck, J. J. (2001) The design of antiparallel coiled coils, *Curr Opin Struct Biol* 11, 450-457.
58. Oakley, M. G., and Kim, P. S. (1998) A buried polar interaction can direct the relative orientation of helices in a coiled coil, *Biochemistry* 37, 12603-12610.

59. Zuccola, H. J., Rozzelle, J. E., Lemon, S. M., Erickson, B. W., and Hogle, J. M. (1998) Structural basis of the oligomerization of hepatitis delta antigen, *Structure* 6, 821-830.
60. Oakley, M. G., and Kim, P. S. (1997) Protein dissection of the antiparallel coiled coil from *Escherichia coli* seryl tRNA synthetase, *Biochemistry* 36, 2544-2549.
61. Zhu, B. Y., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) Packing and hydrophobicity effects on protein folding and stability: effects of beta-branched amino acids, valine and isoleucine, on the formation and stability of two-stranded alpha-helical coiled coils/leucine zippers, *Protein Sci* 2, 383-394.
62. Zhu, H., Celinski, S. A., Scholtz, J. M., and Hu, J. C. (2000) The contribution of buried polar groups to the conformational stability of the GCN4 coiled coil, *Journal of Molecular Biology* 300, 1377-1387.
63. Newman, J. R., and Keating, A. E. (2003) Comprehensive identification of human bZIP interactions with coiled-coil arrays, *Science* 300, 2097-2101.
64. Shen, H., Cao, K., and Wang, X. (2007) A conserved proline residue in the leucine zipper region of AtbZIP34 and AtbZIP61 in *Arabidopsis thaliana* interferes with the formation of homodimer, *Biochem Biophys Res Commun*.
65. Portwich, M., Keller, S., Strauss, H. M., Mahrenholz, C. C., Kretschmar, I., Kramer, A., and Volkmer, R. (2007) A network of coiled-coil associations derived from synthetic GCN4 leucine-zipper arrays, *Angew Chem Int Ed Engl* 46, 1654-1657.
66. Durr, E., Jelesarov, I., and Bosshard, H. R. (1999) Extremely fast folding of a very stable leucine zipper with a strengthened hydrophobic core and lacking electrostatic interactions between helices, *Biochemistry* 38, 870-880.
67. Moran, L. B., Schneider, J. P., Kentsis, A., Reddy, G. A., and Sosnick, T. R. (1999) Transition state heterogeneity in GCN4 coiled coil folding studied by using multisite mutations and crosslinking, *Proc Natl Acad Sci U S A* 96, 10699-10704.
68. Wendt, H., Leder, L., Harma, H., Jelesarov, I., Baici, A., and Bosshard, H. R. (1997) Very rapid, ionic strength-dependent association and folding of a heterodimeric leucine zipper, *Biochemistry* 36, 204-213.
69. Meisner, W. K., and Sosnick, T. R. (2004) Fast folding of a helical protein initiated by the collision of unstructured chains, *Proc Natl Acad Sci U S A* 101, 13478-13482.
70. Myers, J. K., and Oas, T. G. (1999) Reinterpretation of GCN4-p1 folding kinetics: partial helix formation precedes dimerization in coiled coil folding, *J Mol Biol* 289, 205-209.
71. Lee, D. L., Lavigne, P., and Hodges, R. S. (2001) Are trigger sequences essential in the folding of two-stranded alpha-helical coiled-coils?, *J Mol Biol* 306, 539-553.
72. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy, *Biochemistry* 34, 12812-12819.
73. Sosnick, T. R., Jackson, S., Wilk, R. R., Englander, S. W., and DeGrado, W. F. (1996) The role of helix formation in the folding of a fully alpha-helical coiled coil, *Proteins* 24, 427-432.
74. Bosshard, H. R., Durr, E., Hitz, T., and Jelesarov, I. (2001) Energetics of coiled coil folding: the nature of the transition states, *Biochemistry* 40, 3544-3552.
75. Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L., and Matthews, C. R. (2000) Preformed secondary structure drives the association reaction of GCN4-p1, a model coiled-coil system, *J Mol Biol* 296, 1105-1116.
76. Nikolaev, Y., and Pervushin, K. (2007) NMR Spin State Exchange Spectroscopy Reveals Equilibrium of Two Distinct Conformations of Leucine Zipper GCN4 in Solution, *J Am Chem Soc* 129, 6461-6469.
77. d'Avignon, D. A., Bretthorst, G. L., Holtzer, M. E., Schwarz, K. A., Angeletti, R. H., Mints, L., and Holtzer, A. (2006) Site-specific experiments on folding/unfolding of Jun coiled coils: thermodynamic and kinetic parameters from spin inversion transfer nuclear magnetic resonance at leucine-18, *Biopolymers* 83, 255-267.

78. Mason, J. M., Hagemann, U. B., and Arndt, K. M. (2007) Improved stability of the Jun-Fos Activator Protein-1 coiled coil motif: A stopped-flow circular dichroism kinetic analysis, *J Biol Chem* 282, 23015-23024.
79. Karplus, M., and Weaver, D. L. (1994) Protein folding dynamics: the diffusion-collision model and experimental data, *Protein Sci* 3, 650-668.
80. Ibarra-Molero, B., Zitzewitz, J. A., and Matthews, C. R. (2004) Salt-bridges can stabilize but do not accelerate the folding of the homodimeric coiled-coil peptide GCN4-p1, *J Mol Biol* 336, 989-996.
81. Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., and Hallewell, R. A. (1992) Faster superoxide dismutase mutants designed by enhancing electrostatic guidance, *Nature* 358, 347-351.
82. Camacho, C. J., Kimura, S. R., DeLisi, C., and Vajda, S. (2000) Kinetics of desolvation-mediated protein-protein binding, *Biophys J* 78, 1094-1105.
83. Huang, C. Y., Getahun, Z., Zhu, Y., Klemke, J. W., DeGrado, W. F., and Gai, F. (2002) Helix formation via conformation diffusion search, *Proc Natl Acad Sci U S A* 99, 2788-2793.
84. Hummer, G., Garcia, A. E., and Garde, S. (2000) Conformational diffusion and helix formation kinetics, *Phys Rev Lett* 85, 2637-2640.
85. Thompson, P. A., Eaton, W. A., and Hofrichter, J. (1997) Laser temperature jump study of the helix $\rightleftharpoons$ coil kinetics of an alanine peptide interpreted with a 'kinetic zipper' model, *Biochemistry* 36, 9200-9210.
86. Williams, S., Causgrove, T. P., Gilmanishin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1996) Fast events in protein folding: helix melting and formation in a small peptide, *Biochemistry* 35, 691-697.
87. Holtzer, M. E., Bretthorst, G. L., d'Avignon, D. A., Angeletti, R. H., Mints, L., and Holtzer, A. (2001) Temperature dependence of the folding and unfolding kinetics of the GCN4 leucine zipper via <sup>13</sup>C(alpha)-NMR, *Biophys J* 80, 939-951.
88. d'Avignon, D. A., Bretthorst, G. L., Holtzer, M. E., and Holtzer, A. (1999) Thermodynamics and kinetics of a folded-folded' transition at valine-9 of a GCN4-like leucine zipper, *Biophys J* 76, 2752-2759.
89. Dragan, A. I., and Privalov, P. L. (2002) Unfolding of a leucine zipper is not a simple two-state transition, *J Mol Biol* 321, 891-908.
90. Knappenberger, J. A., Smith, J. E., Thorpe, S. H., Zitzewitz, J. A., and Matthews, C. R. (2002) A buried polar residue in the hydrophobic interface of the coiled-coil peptide, GCN4-p1, plays a thermodynamic, not a kinetic role in folding, *J Mol Biol* 321, 1-6.
91. Karplus, M., and Weaver, D. L. (1976) Protein-folding dynamics, *Nature* 260, 404-406.
92. Kammerer, R. A., Jaravine, V. A., Frank, S., Schulthess, T., Landwehr, R., Lustig, A., Garcia-Echeverria, C., Alexandrescu, A. T., Engel, J., and Steinmetz, M. O. (2001) An intrahelical salt bridge within the trigger site stabilizes the GCN4 leucine zipper, *J Biol Chem* 276, 13685-13688.
93. Saudek, V., Pastore, A., Morelli, M. A., Frank, R., Gausepohl, H., and Gibson, T. (1991) The solution structure of a leucine-zipper motif peptide, *Protein Eng* 4, 519-529.
94. Matousek, W. M., Ciani, B., Fitch, C. A., Garcia-Moreno, B., Kammerer, R. A., and Alexandrescu, A. T. (2007) Electrostatic contributions to the stability of the GCN4 leucine zipper structure, *J Mol Biol* 374, 206-219.
95. Wendt, H., Berger, C., Baici, A., Thomas, R. M., and Bosshard, H. R. (1995) Kinetics of folding of leucine zipper domains, *Biochemistry* 34, 4097-4107.
96. Lovett, E. G., D'Avignon, D. A., Holtzer, M. E., Braswell, E. H., Zhu, D., and Holtzer, A. (1996) Observation via one-dimensional <sup>13</sup>Calpha NMR of local conformational substates in thermal unfolding equilibria of a synthetic analog of the GCN4 leucine zipper, *Proc Natl Acad Sci U S A* 93, 1781-1785.

97. Holtzer, M. E., Lovett, E. G., d'Avignon, D. A., and Holtzer, A. (1997) Thermal unfolding in a GCN4-like leucine zipper:  $^{13}\text{C}$  alpha NMR chemical shifts and local unfolding curves, *Biophys J* 73, 1031-1041.
98. d'Avignon, D. A., Bretthorst, G. L., Holtzer, M. E., and Holtzer, A. (1998) Site-specific thermodynamics and kinetics of a coiled-coil transition by spin inversion transfer NMR, *Biophys J* 74, 3190-3197.
99. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1997) Salt effects on protein stability: two-stranded alpha-helical coiled-coils containing inter- or intrahelical ion pairs, *J Mol Biol* 267, 1039-1052.
100. Hendsch, Z. S., and Tidor, B. (1999) Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding, *Protein Sci* 8, 1381-1392.
101. Junius, F. K., O'Donoghue, S. I., Nilges, M., Weiss, A. S., and King, G. F. (1996) High resolution NMR solution structure of the leucine zipper domain of the c-Jun homodimer, *J Biol Chem* 271, 13663-13667.
102. Patel, L. R., Curran, T., and Kerppola, T. K. (1994) Energy transfer analysis of Fos-Jun dimerization and DNA binding, *Proc Natl Acad Sci U S A* 91, 7360-7364.
103. Cranz, S., Berger, C., Baici, A., Jelesarov, I., and Bosshard, H. R. (2004) Monomeric and dimeric bZIP transcription factor GCN4 bind at the same rate to their target DNA site, *Biochemistry* 43, 718-727.
104. Pace, C. N., and Scholtz, J. M. (1998) A helix propensity scale based on experimental studies of peptides and proteins, *Biophys J* 75, 422-427.
105. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., and Lassar, A. B. (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence, *Cell* 58, 537-544.
106. Ruberti, I., Sessa, G., Lucchetti, S., and Morelli, G. (1991) A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif, *Embo J* 10, 1787-1791.
107. Tupler, R., Perini, G., and Green, M. R. (2001) Expressing the human genome, *Nature* 409, 832-833.
108. Ledent, V., Paquet, O., and Vervoort, M. (2002) Phylogenetic analysis of the human basic helix-loop-helix proteins, *Genome Biol* 3, RESEARCH0030.
109. Eferl, R., and Wagner, E. F. (2003) AP-1: a double-edged sword in tumorigenesis, *Nat Rev Cancer* 3, 859-868.
110. Correa, L. G., Riano-Pachon, D. M., Schrago, C. G., dos Santos, R. V., Mueller-Roeber, B., and Vincentz, M. (2008) The role of bZIP transcription factors in green plant evolution: adaptive features emerging from four founder genes, *PLoS ONE* 3, e2944.
111. Massari, M. E., and Murre, C. (2000) Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms, *Mol Cell Biol* 20, 429-440.
112. Luscher, B. (2001) Function and regulation of the transcription factors of the Myc/Max/Mad network, *Gene* 277, 1-14.
113. Meyer, N., and Penn, L. Z. (2008) Reflecting on 25 years with MYC, *Nat Rev Cancer* 8, 976-990.
114. Patel, J. H., Loboda, A. P., Showe, M. K., Showe, L. C., and McMahon, S. B. (2004) Analysis of genomic targets reveals complex functions of MYC, *Nat Rev Cancer* 4, 562-568.
115. Ariel, F. D., Manavella, P. A., Dezar, C. A., and Chan, R. L. (2007) The true story of the HD-Zip family, *Trends Plant Sci* 12, 419-426.
116. Baranger, A. M. (1998) Accessory factor-bZIP-DNA interactions, *Curr Opin Chem Biol* 2, 18-23.

117. Boer, U., Eglins, J., Krause, D., Schnell, S., Schofl, C., and Knepel, W. (2007) Enhancement by lithium of cAMP-induced CREB/CREB-directed gene transcription conferred by TORC on the CREB basic leucine zipper domain, *Biochem J* 408, 69-77.
118. Waldmann, I., Walde, S., and Kehlenbach, R. H. (2007) Nuclear import of c-Jun is mediated by multiple transport receptors, *J Biol Chem* 282, 27685-27692.
119. Bailey, D., and O'Hare, P. (2007) Transmembrane bZIP transcription factors in ER stress signaling and the unfolded protein response, *Antioxid Redox Signal* 9, 2305-2321.
120. Tajima, H., Iwata, Y., Iwano, M., Takayama, S., and Koizumi, N. (2008) Identification of an Arabidopsis transmembrane bZIP transcription factor involved in the endoplasmic reticulum stress response, *Biochem Biophys Res Commun* 374, 242-247.
121. Shifera, A. S., Friedman, J. M., and Horwitz, M. S. (2008) IKKgamma (NEMO) is involved in the coordination of the AP-1 and NF-kappaB pathways, *Mol Cell Biochem* 310, 181-190.
122. Takemori, H., Kajimura, J., and Okamoto, M. (2007) TORC-SIK cascade regulates CREB activity through the basic leucine zipper domain, *Febs J* 274, 3202-3209.
123. Karanam, B., Wang, L., Wang, D., Liu, X., Marmorstein, R., Cotter, R., and Cole, P. A. (2007) Multiple roles for acetylation in the interaction of p300 HAT with ATF-2, *Biochemistry* 46, 8207-8216.
124. Zhou, G. K., Kubo, M., Zhong, R., Demura, T., and Ye, Z. H. (2007) Overexpression of miR165 affects apical meristem formation, organ polarity establishment and vascular development in Arabidopsis, *Plant Cell Physiol* 48, 391-404.
125. Schutze, K., Harter, K., and Chaban, C. (2008) Post-translational regulation of plant bZIP factors, *Trends Plant Sci* 13, 247-255.
126. Norton, J. D. (2000) ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis, *Journal of cell science* 113 ( Pt 22), 3897-3905.
127. Ruzinova, M. B., and Benezra, R. (2003) Id proteins in development, cell cycle and cancer, *Trends in cell biology* 13, 410-418.
128. Wenkel, S., Emery, J., Hou, B. H., Evans, M. M., and Barton, M. K. (2007) A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes, *The Plant cell* 19, 3379-3390.
129. Kiessling, A., Sperl, B., Hollis, A., Eick, D., and Berg, T. (2006) Selective inhibition of c-Myc/Max dimerization and DNA binding by small molecules, *Chem Biol* 13, 745-751.
130. Berg, T. (2008) Inhibition of transcription factors with small organic molecules, *Curr Opin Chem Biol* 12, 464-471.
131. Aikawa, Y., Morimoto, K., Yamamoto, T., Chaki, H., Hashiramoto, A., Narita, H., Hirono, S., and Shiozawa, S. (2008) Treatment of arthritis with a selective inhibitor of c-Fos/activator protein-1, *Nat Biotechnol* 26, 817-823.
132. Zhang, J. W., Tang, Q. Q., Vinson, C., and Lane, M. D. (2004) Dominant-negative C/EBP disrupts mitotic clonal expansion and differentiation of 3T3-L1 preadipocytes, *Proc Natl Acad Sci U S A* 101, 43-47.
133. Gerdes, M. J., Myakishev, M., Frost, N. A., Rishi, V., Moitra, J., Acharya, A., Levy, M. R., Park, S. W., Glick, A., Yuspa, S. H., and Vinson, C. (2006) Activator protein-1 activity regulates epithelial tumor cell identity, *Cancer research* 66, 7578-7588.
134. Strauss, H. M., and Keller, S. (2008) Pharmacological interference with protein-protein interactions mediated by coiled-coil motifs, *Handbook of experimental pharmacology*, 461-482.
135. Li, X., and Stark, G. R. (2002) NFkappaB-dependent signaling pathways, *Exp Hematol* 30, 285-296.
136. DiDonato, J. A., Hayakawa, M., Rothwarf, D. M., Zandi, E., and Karin, M. (1997) A cytokine-responsive IkappaB kinase that activates the transcription factor NF-kappaB, *Nature* 388, 548-554.

137. Zandi, E., Rothwarf, D. M., Delhase, M., Hayakawa, M., and Karin, M. (1997) The IkappaB kinase complex (IKK) contains two kinase subunits, IKKalpha and IKKbeta, necessary for IkappaB phosphorylation and NF-kappaB activation, *Cell* 91, 243-252.
138. Rothwarf, D. M., Zandi, E., Natoli, G., and Karin, M. (1998) IKK-gamma is an essential regulatory subunit of the IkappaB kinase complex, *Nature* 395, 297-300.
139. Karin, M. (1999) How NF-kappaB is activated: the role of the IkappaB kinase (IKK) complex, *Oncogene* 18, 6867-6874.
140. Hacker, H., and Karin, M. (2006) Regulation and function of IKK and IKK-related kinases, *Sci STKE* 2006, re13.
141. Zandi, E., Chen, Y., and Karin, M. (1998) Direct phosphorylation of IkappaB by IKKalpha and IKKbeta: discrimination between free and NF-kappaB-bound substrate, *Science* 281, 1360-1363.
142. Park, M. Y., Moon, J. H., Lee, K. S., Choi, H. I., Chung, J., Hong, H. J., and Kim, E. (2007) FAF1 suppresses IkappaB kinase (IKK) activation by disrupting the IKK complex assembly, *J Biol Chem* 282, 27572-27577.
143. Bornberg-Bauer, E., Rivals, E., and Vingron, M. (1998) Computational approaches to identify leucine zippers, *Nucleic Acids Res* 26, 2740-2746.
144. Wolf, E., Kim, P. S., and Berger, B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils, *Protein Sci* 6, 1179-1189.
145. Di Marco, B., Massetti, M., Bruscoli, S., Macchiarulo, A., Di Virgilio, R., Velardi, E., Donato, V., Migliorati, G., and Riccardi, C. (2007) Glucocorticoid-induced leucine zipper (GILZ)/NF-kappaB interaction: role of GILZ homo-dimerization and C-terminal domain, *Nucleic Acids Res* 35, 517-528.
146. Surks, H. K., Mochizuki, N., Kasai, Y., Georgescu, S. P., Tang, K. M., Ito, M., Lincoln, T. M., and Mendelsohn, M. E. (1999) Regulation of myosin phosphatase by a specific interaction with cGMP- dependent protein kinase Ialpha, *Science* 286, 1583-1587.
147. Lee, E., Hayes, D. B., Langsetmo, K., Sundberg, E. J., and Tao, T. C. (2007) Interactions between the leucine-zipper motif of cGMP-dependent protein kinase and the C-terminal region of the targeting subunit of myosin light chain phosphatase, *J Mol Biol* 373, 1198-1212.
148. Ihara, E., and MacDonald, J. A. (2007) The regulation of smooth muscle contractility by zipper-interacting protein kinase, *Can J Physiol Pharmacol* 85, 79-87.
149. Kawai, T., Matsumoto, M., Takeda, K., Sanjo, H., and Akira, S. (1998) ZIP kinase, a novel serine/threonine kinase which mediates apoptosis, *Mol Cell Biol* 18, 1642-1651.
150. Graves, P. R., Winkfield, K. M., and Haystead, T. A. (2005) Regulation of zipper-interacting protein kinase activity in vitro and in vivo by multisite phosphorylation, *J Biol Chem* 280, 9363-9374.
151. Hulme, J. T., Ahn, M., Hauschka, S. D., Scheuer, T., and Catterall, W. A. (2002) A novel leucine zipper targets AKAP15 and cyclic AMP-dependent protein kinase to the C terminus of the skeletal muscle Ca2+ channel and modulates its function, *J Biol Chem* 277, 4079-4087.
152. Hulme, J. T., Lin, T. W., Westenbroek, R. E., Scheuer, T., and Catterall, W. A. (2003) Beta-adrenergic regulation requires direct anchoring of PKA to cardiac CaV1.2 channels via a leucine zipper interaction with A kinase-anchoring protein 15, *Proc Natl Acad Sci U S A* 100, 13093-13098.
153. Few, W. P., Scheuer, T., and Catterall, W. A. (2007) Dopamine modulation of neuronal Na(+) channels requires binding of A kinase-anchoring protein 15 and PKA by a modified leucine zipper motif, *Proc Natl Acad Sci U S A* 104, 5187-5192.
154. Sollner, T., Whiteheart, S. W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J. E. (1993) SNAP receptors implicated in vesicle targeting and fusion, *Nature* 362, 318-324.

155. Fasshauer, D., Sutton, R. B., Brunger, A. T., and Jahn, R. (1998) Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs, *Proc Natl Acad Sci U S A* 95, 15781-15786.
156. Sutton, R. B., Fasshauer, D., Jahn, R., and Brunger, A. T. (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution, *Nature* 395, 347-353.
157. Mobley, P. W., Barry, J. A., Waring, A. J., Sherman, M. A., and Gordon, L. M. (2007) Membrane perturbing actions of HIV type 1 glycoprotein 41 domains are inhibited by helical C-peptides, *AIDS Res Hum Retroviruses* 23, 224-242.
158. Knežlik, Z., Smekalova, Z., Ruml, T., and Sakalian, M. (2007) Multimerization of the p12 domain is necessary for Mason-Pfizer monkey virus Gag assembly in vitro, *Virology* 365, 260-270.
159. Nellissery, J. K., Szczepaniak, R., Lamberti, C., and Weller, S. K. (2007) A putative leucine zipper within the herpes simplex virus type 1 UL6 protein is required for portal ring formation, *J Virol* 81, 8868-8877.
160. Haller, O., Staeheli, P., and Kochs, G. (2007) Interferon-induced Mx proteins in antiviral host defense, *Biochimie* 89, 812-818.
161. Melen, K., Ronni, T., Broni, B., Krug, R. M., von Bonsdorff, C. H., and Julkunen, I. (1992) Interferon-induced Mx proteins form oligomers and contain a putative leucine zipper, *J Biol Chem* 267, 25898-25907.
162. Haller, O., Frese, M., and Kochs, G. (1998) Mx proteins: mediators of innate resistance to RNA viruses, *Revue scientifique et technique (International Office of Epizootics)* 17, 220-230.
163. Tumpey, T. M., Szretter, K. J., Van Hoeven, N., Katz, J. M., Kochs, G., Haller, O., Garcia-Sastre, A., and Staeheli, P. (2007) The Mx1 gene protects mice against the pandemic 1918 and highly lethal human H5N1 influenza viruses, *J Virol* 81, 10818-10821.
164. Haller, O., and Kochs, G. (2002) Interferon-induced mx proteins: dynamin-like GTPases with antiviral activity, *Traffic* 3, 710-717.
165. Di Paolo, C., Hefti, H. P., Meli, M., Landis, H., and Pavlovic, J. (1999) Intramolecular backfolding of the carboxyl-terminal end of MxA protein is a prerequisite for its oligomerization, *J Biol Chem* 274, 32071-32078.
166. Janzen, C., Kochs, G., and Haller, O. (2000) A monomeric GTPase-negative MxA mutant with antiviral activity, *J Virol* 74, 8202-8206.
167. Schumacher, B., and Staeheli, P. (1998) Domains mediating intramolecular folding and oligomerization of MxA GTPase, *J Biol Chem* 273, 28365-28370.
168. Zurcher, T., Pavlovic, J., and Staeheli, P. (1992) Mechanism of human MxA protein action: variants with changed antiviral properties, *Embo J* 11, 1657-1661.
169. Kochs, G., Janzen, C., Hohenberg, H., and Haller, O. (2002) Antivirally active MxA protein sequesters La Crosse virus nucleocapsid protein into perinuclear complexes, *Proc Natl Acad Sci U S A* 99, 3153-3158.
170. Andersson, I., Bladh, L., Mousavi-Jazi, M., Magnusson, K. E., Lundkvist, A., Haller, O., and Mirazimi, A. (2004) Human MxA protein inhibits the replication of Crimean-Congo hemorrhagic fever virus, *J Virol* 78, 4323-4329.