Published in final edited form as:
J Mol Biol. 2011 April 29; 408(2): 304–320. doi:10.1016/j.jmb.2011.02.046.

Design of peptide inhibitors that bind the bZIP domain of Epstein-Barr virus protein BZLF1

T. Scott Chen, **Aaron W. Reinke**, and **Amy E. Keating**
Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139

Abstract

Designing proteins or peptides that bind native protein targets can aid the development of novel reagents and/or therapeutics. Rational design also tests our understanding of the principles underlying protein recognition. This paper describes several strategies used to design peptides to bind to the bZIP domain of the viral transcription factor BZLF1, which is encoded by the Epstein-Barr virus (EBV). BZLF1 regulates transition of EBV from a latent to a lytic state. It shares some properties in common with more-studied human bZIP transcription factors, but also includes novel structural elements that pose interesting challenges to inhibitor design. In designing peptides that bind to BZLF1 by forming a coiled-coil structure, we considered both affinity for BZLF1 and also undesired self-association, which can weaken the effectiveness of an inhibitor. Several designed peptides exhibited different degrees of target-binding affinity and self-association. Rationally engineered molecules were more potent inhibitors of DNA binding than a control peptide corresponding to the native BZLF1 dimerzation region itself. The most potent inhibitors included both positive and negative design elements and exploited interaction with the coiled-coil and basic DNA-binding regions of BZLF1.

Keywords

BZLF1; protein design; protein-protein interaction inhibitor; bZIP transcription factors; interaction specificity

Introduction

The basic-region leucine-zipper (bZIP) transcription factors are a large class of proteins conserved in eukaryotes and several viruses that regulate a wide range of biological processes. The structure of bZIP-DNA complexes is very simple: a helical and positively charged DNA-binding region is contiguous with a coiled coil that mediates protein homo- or hetero-dimerization.[1] The bZIP coiled-coil helices wrap around one another in a parallel orientation with “knobs-into-holes” side-chain packing geometry, and a 7-amino-acid heptad repeat characterizes the structure, in which each residue can be assigned a register position labeled *a* through *g* (Fig. 1). High-affinity binding of bZIP transcription factors to DNA requires protein dimerization.

Corresponding author: Amy E. Keating, 77 Massachusetts Avenue, Building 68-622, Cambridge, Massachusetts 02139, USA, Telephone: (617) 452-3398, Fax: (617) 253-4043, keating@mit.edu.

Given the many important biological roles of the bZIPs, molecules that selectively disrupt bZIP-DNA interactions could be valuable reagents and even potential therapeutics. Several strategies have been reported for identifying inhibitors. Small molecules have been discovered via high-throughput screening,[2, 3] and peptides that bind to the coiled-coil regions of the bZIPs and disrupt dimer formation have been selected from targeted combinatorial libraries.[4, 5, 6] A particularly effective strategy for blocking bZIP-DNA interactions was developed by Vinson and co-workers, who created a series of dominant-negative peptide inhibitors by replacing the basic regions of certain bZIP proteins with a sequence enriched in negatively charged residues (the “acidic extension”), giving so-called A-ZIPs.[7, 8, 9, 10] The A-ZIPs bind tightly and selectively to bZIPs and have been used to study the effects of inhibiting dimerization and hence DNA binding in both cell culture and animal models.[11, 12]

Current understanding of bZIP coiled-coil interactions has also enabled the computational design of synthetic peptides to block bZIP dimerization. Significant effort has been dedicated to elucidating sequence determinants governing the interactions of bZIP coiled coils, and to developing predictive computational models that capture these. Several types of residue-pair interactions that are important for specificity have been characterized in detail over the past 20 years, and models derived from physics-based calculations, machine learning, and experimentally measured coupling energies have been developed to explain and predict bZIP coiled-coil interactions.[4, 13, 14, 15, 16, 17] Using such binding models, Grigoryan et al. recently designed a series of peptides that bind to targets in 19 out of 20 human bZIP families.[18]

An interesting issue in the study of bZIP interactions is specificity. Given the similarities among sequences, and the many bZIPs in most eukaryotes, a large number of homo- and heterodimers can potentially form. Interactions among human bZIPs have been shown to be highly selective when assayed *in vitro*,[19, 20] but it can be difficult to achieve specificity in designed bZIP-like peptides. In particular, peptides engineered to bind to bZIP coiled-coil regions have been shown to self-associate strongly and also interact with undesired partners.[5, 18] In this work we address considerations of both affinity and anti-homodimer specificity in the design of peptide inhibitors for a viral bZIP protein, BZLF1.

BZLF1 (Zta, ZEBRA, EB1) is encoded by the Epstein-Barr virus (EBV) and triggers the virus’s latent to lytic switch by functioning as a transcription factor and regulator of DNA replication.[21, 22, 23, 24] Infection by EBV has been linked to several human malignancies such as Hodgkin’s disease and Burkitt’s lymphoma.[25] The basic region of BZLF1 is highly homologous to that of human bZIPs and is responsible for direct contact with DNA; a coiled-coil region immediately C-terminal to the basic helix mediates dimerization. However, a recent crystal structure and other biochemical studies have revealed several unique features of BZLF1 (Fig. 1a).[26, 27] The coiled-coil region at the dimerization interface is only 4 heptads long, whereas the coiled-coil regions of human bZIPs typically contain at least 5 heptads. Furthermore, only one of the four BZLF1 coiled-coil heptads includes a leucine residue at the *d* position; this residue occurs with much higher frequency in human bZIP sequences (hence the name “leucine zipper”). The stability of the BZLF1 homodimer is significantly enhanced by a unique C-terminal (CT) region that folds back on the coiled coil to form additional contacts;[27] the CT region is only partially observed in the crystal structure. Prior work using peptide arrays showed that BZLF1 constructs corresponding to the coiled coil or the coiled coil plus the CT region homo-associate in preference to binding any of 33 representative human bZIP proteins.[28]

It has been shown that a peptide corresponding to the coiled-coil region of BZLF1, lacking the DNA binding residues, inhibits BZLF1 binding to DNA at high micromolar

concentrations.[29] In this work, we sought new peptides that would mimic the coiled-coil interface of the native structure yet provide more potent inhibition of DNA binding. As a design target, BZLF1 is both simpler and more complex than human and viral bZIPs that have been the subjects of previous computational design studies.[18, 28] It is simpler because of its unique structural features, which make coiled-coil inhibitors designed to target it unlikely to interact broadly with other bZIP proteins. However, it is more complex because the CT region and unusually tight helix packing make the interface unlike the dimerization domains of better-understood bZIPs.[26] Here we explore the extent to which previously applied design strategies can be used successfully in the context of BZLF1. Throughout our analyses, we explicitly addressed two design criteria: affinity for BZLF1 and design self-association, which is an undesirable trait for an inhibitor. The best inhibitor incorporated both elements and included modifications of BZLF1 in both the coiled- coil and DNA-binding regions. As assessed using DNA-binding gel-shift assays, this designed peptide was much more potent than one corresponding to the native dimerization domain.

Results

Computational design of a peptide to bind the N-terminal part of the BZLF1 coiled coil

Our goal was to identify variants of the BZLF1 dimerization domain that would function as more effective dominant negative inhibitors of DNA binding. As described in the Introduction, BZLF1 possesses several unique features as a bZIP design target. These include the unconventional, short coiled-coil segment and the CT region. The CT can be divided into the proximal CT (residues 222–231) and the less structured distal CT (residues 232–246), as shown in Fig. 1b. We began by re-designing the N-terminal two and a half heptads of the BZLF1 coiled coil (residues 191–209, Fig 2b), because we anticipated that this segment would provide the greatest opportunity to improve affinity and heterodimer specificity over the native sequence. Residues 210–221 also form part of the coiled-coil structure, but additionally engage in non-coiled-coil hydrophobic contacts with the proximal CT as observed in the crystal structure (Fig. 1a). Thus, in order to maintain this stabilizing interaction, these residues were not changed in the design.

Both the desired design-target heterodimer and the undesired design homodimer were modeled as parallel, blunt ended coiled coils. We used the CLASSY protein-design algorithm to choose residues at 10 sites in the design, optimizing the predicted affinity of the design-target complex.[18] The scoring function used was based on a hybrid model that included both physics-based and experimentally derived terms and is described further in the Methods. The optimal-affinity design, which we call BD_{CC} (BZLF1 design against the coiled-coil region, shown in Fig. 2c), was predicted to be hetero-specific. In design energy units the predicted stabilities were as follows: BZLF1 homodimer: −29 kcal/mol, BD_{CC} homodimer: −32 kcal/mol, BZLF1/BD_{CC} heterodimer: −44 kcal/mol. Although the score for the design self-interaction was close to that for native BZLF1 coiled-coil homodimerization, the score for the design-target interaction was significantly better. Thus, although CLASSY can be used to improve specificity against undesired states as well as affinity for a target,[18] this was predicted not to be necessary in this case.

The BD_{CC} solution populated most *a* and *d* positions (coiled-coil “core” positions) with Ile and Leu respectively, which are very common in conventional bZIP sequences (Fig. 2c). A single *d*-position Glu residue at the extreme N terminus of the coiled coil is uncharacteristic of bZIP sequences, but was predicted to interact favorably with an *e*-position Lys on BZLF1. The five designed *e* and *g* positions (coiled-coil “edge” positions) were all populated with glutamate for improved electrostatic interactions with the target, where three residues in this region are positively charged. Interestingly, predicted charged interactions involved both edge-to-edge (e.g. *g* to *e*′) and core-to-edge (*d* to *e*′) residues in the BZLF1

target, as was previously observed for anti-human bZIP designs.[18] Although core sites occupied by Ile and Leu favor design self-interaction, the charged residues at *e* and *g* are predicted to disfavor it. Charge repulsion is a commonly observed negative design element in many native and model coiled coils.[30, 31, 32, 33]

The anti-BZLF1 peptide was cloned in the context of residues 191–231 of BZLF1. This construct, BD_{CC}^{231}, includes the entire coiled-coil domain and the proximal CT (Fig. 1b, 2a, Table 1), potentially retaining native interactions observed in the X-ray structure between the C-terminal part of the coiled coil and the CT region. Because the residues optimized in the design calculations are more than 8 Å away from residue I231 in the modeled structure (Fig. 2a), the proximal CT excluded from the calculations was not expected to significantly influence the results. Potential interactions between the designed residues and the distal CT, which are not evident in the crystal structure but are suggested by prior studies[27], are addressed in experiments described below.

We used circular dichroism (CD) spectroscopy to study the interaction properties of BD_{CC}^{231}. Thermal denaturation experiments showed that the BD_{CC}^{231} homo-oligomer is destabilized compared to the target homodimer in the same sequence context ($BZLF1^{231}$, residues 191 to 231); T_m values were 38 °C vs. 43 °C (Fig. 3a and Table 1). The hetero-complex between BD_{CC}^{231} and $BZLF1^{231}$ (T_m of 53 °C, Table 2) was significantly stabilized compared to the $BZLF1^{231}$ homodimer. We conclude that the BD_{CC}^{231} design is very hetero-specific, consistent with expectations based on the design algorithm. The agreement indicates success of the automated CLASSY approach even on a target with a sequence quite different from the human bZIPs.

Designs with weaker self-association

The BD_{cc} design achieved hetero-specificity mostly by improving design-target affinity compared to the native BZLF1 complex. We also sought solutions that achieved hetero-specificity against the same target (the N-terminal part of the BZLF1 coiled coil) by weakening design self-interaction. Toward this end, we tested a negative design strategy that placed charged residues at a core *d* position and the adjacent *e* position such that they would create a local cluster of 4 negative charges in the modeled design coiled-coil homodimer. There are 3 close inter-chain pair contacts in such a cluster (2 *d*–*e′* interactions and one *d*-*d′* interaction). We observed variations of this strategy in design solutions obtained using the CLASSY algorithm when optimizing affinity for the target under increasingly stringent constraints limiting the stability of the design homodimer.

We picked two sets of amino-acid changes, (K207E, S208D) and (Y200E, R201E), each corresponding to the (*d*, *e′*) negative design strategy described above. We also included one stabilizing design element present in the BD_{CC}^{231} solution, A204I (substituting lie for Ala at an *a* position), to compensate for a potential loss in stability due to the introduction of charge in the core. The resulting two designs were cloned, expressed and purified, again in the context of BZLF1 residues 191 to 231 (204I, 207E, 208D, referred to as BD_{IED}^{231}, and 200E, 201E, 204I, referred to as BD_{EEI}^{231}, Fig 2d–e).

Thermal denaturation experiments monitored by CD showed that both designed peptides, BD_{IED}^{231} and BD_{EEI}^{231}, had relatively weak helical signals even at very low temperatures (Fig. 3b, c), illustrating the effectiveness of the negative design strategy. We compared the melting curve for the mixture of each design and $BZLF1^{231}$ with the numerical average of the

individual melting curves for each species (Fig. 3b, c). The difference between the two curves below ~22 °C reflects interaction between the designed peptides and $BZLF1^{231}$, and confirms that the designed peptides bind the target more strongly than they interact with themselves. However, an interaction is evident only at low temperatures, indicating that the stability of the design-target complex is lower than the $BZLF1^{231}$ target homodimer. Therefore, these 2 designed peptides represent a specificity profile distinct from that of BD_{CC}^{231}; one that achieves greater destabilization against design self-interaction at the expense of the stability of the design-target interaction.

BD_{CC} and BZLF1 form a heterodimer

We modeled all coiled-coil interactions as parallel, symmetric dimers. Although the oligomerization states of coiled coils can be sensitive to very few amino-acid changes,[34, 35] in BZLF1 the presence of the CT region is expected to strongly favor the parallel dimer geometry observed in the crystal structure for BZLF1. The designed heterodimer also includes an Asn-Asn interaction at *a*-*a′*, which has been shown to strongly favor dimers, and multiple charged residues at the *e* and *g* positions that are also more prevalent in dimers.[36] Nevertheless, we performed analytical ultracentrifugation (AUC) experiments to study the interaction between BD_{CC}^{231} and $BZLF1^{231}$. Global analysis of sedimentation equilibrium runs performed at multiple concentrations and rotor speeds showed that the best-fit molecular weight for both BD_{CC}^{231} and the 1:1 mixture of BD_{CC}^{231} with $BZLF1^{231}$ corresponded to that expected for a dimer (representative data are shown along with the global fit in Fig. 4). For BD_{CC}^{231} with $BZLF1^{231}$, the fitted molecular weight was 104% of that expected for the heterodimer, with a fitted RMS of 0.027 fringes. RMS values obtained by fixing an exact dimer or trimer weight were 0.029 or 0.090 fringes, respectively. For BD_{CC}^{231}, the fitted molecular weight is 102% of that expected for the homodimer, with a fitted RMS of 0.021 fringes. RMS values obtained by fixing a dimer or a trimer weight were 0.021 or 0.10 fringes, respectively. The AUC data thus confirm the validity of modeling these interactions as dimers.

Testing designs in the full-length BZLF1 dimerization domain

The designs described above targeted the BZLF1 coiled coil and were tested in the context of $BZLF1^{231}$. However, inhibitors of protein function must bind to the full-length protein. One difficulty with designing against the entire BZLF1 dimerization domain (residues 191–245) is that the crystal structure shows only the proximal and part of the distal CT region (up to residue 236), with the remaining part of the distal CT region contributing no electron density.[26] Nevertheless, the distal CT region (Fig. 1b) has been shown to contribute positively to BZLF1 dimer stability despite possibly being less structured.[27]

We tested whether our design procedures, which considered only the structured coiled coil, could provide molecules that bind the full-length BZLF1 dimerization domain. For this purpose, a BZLF1 construct that included both the DNA binding basic region and the full-length dimerization domain (termed B-$BZLF1^{245}$, residues 175–245, Table 1) was used instead of $BZLF1^{231}$ as the target. The designed mutations in BD_{CC}^{231} and BD_{IED}^{231}were made in the context of the full-length BZLF1 dimerization domain without the basic region (residues 191–245) to create two new design constructs, BD_{CC}^{245} and BD_{IED}^{245} (Fig. 2a, Table 1); the distal CT was included in the design constructs to exploit its potentially favorable interaction with the target.

The distal CT dramatically stabilized the BZLF1 homodimer (compare $BZLF1^{231}$ and $BZLF1^{245}$ T_m values of 43 °C and 71 °C, respectively), consistent with prior reports.[27] In

contrast, self-association of the BD_{CC} design was not significantly stabilized by the distal CT (Table 1). When BD_{CC}^{245} and B-$BZLF1^{245}$ were mixed, there was clear evidence of interaction (Fig. 3d, Table 2). However, the hetero-interaction between BD_{CC}^{245} and B-$BZLF1^{245}$ did not appear to be stronger than the self-association of the target B-$BZLF1^{245}$ (Table 1, 2), which contrasts with the behavior of the shorter constructs, BD_{CC}^{231} and $BZLF1^{231}$ (Fig. 3a, Table 2). Differences in relative stabilities for the shorter and longer constructs suggest that residues in the design do not interact as favorably as the native residues with the distal CT.

In contrast to BD_{CC}^{245}, analysis of BD_{IED}^{245} showed that both the design self-interaction and the design-target interaction were stabilized by the distal CT (compare Fig. 3b with Fig. 3e). As a result, BD_{IED}^{245} was heterospecific at low temperature. Compared to BD_{CC}^{245}, BD_{IED}^{245} showed weaker self-association but also displayed weaker affinity for B-$BZLF1^{245}$. Together, the results show that the effect of the distal CT is not negligible and depends on sequence in the coiled-coil region. The impact of the distal CT on the specificity profiles for different designs is considered further in the Discussion.

Specificity of BD_{CC} against human bZIPs

Specificity against human bZIP proteins was not addressed explicitly in our design procedure because we reasoned that the CT region, which is unique to BZLF1, would likely stabilize interaction with BZLF1 but not with human proteins. To assess this, we selected a few human bZIPs and evaluated their interactions with BD_{CC} using CD spectroscopy. To identify those human bZIP proteins most likely to associate with BD_{CC}, we calculated interaction scores with 36 representative human bZIP coiled coils using the scoring function employed in the CLASSY algorithm, which has been shown to be useful for evaluating bZIP coiled-coil associations (Fig. 5a).[13, 18] Interestingly, BD_{CC} was predicted to interact more favorably with BZLF1 than with any of the human bZIPs, even though the scoring scheme used did not consider interactions involving the CT region. We chose 5 of the top 10 scoring complexes for experimental testing, selecting representative proteins that spanned 5 families and included JUN, the closest predicted competitor. We used constructs for the human proteins that included the basic region and the coiled coil (Fig. 5b–f). Analysis of melting curves for each human bZIP and each 1:1 mixture with BD_{CC}^{245} showed that only JUN interacted with BD_{CC}^{245}. The BD_{CC}^{245}/JUN complex, however, was significantly weaker than that between BD_{CC}^{245} and B-$BZLF1^{245}$ (T_m values of 41°C vs. 66°C, Table 2). Thus, BD_{CC} is not a promiscuous design and binds preferentially to its target, BZLF1.

Enhancing design performance with an N-terminal acidic extension

Vinson and colleagues have shown that replacing the basic region of several native bZIPs with a designed sequence enriched in glutamates can provide potent dominant-negative inhibitors of bZIP dimerization and DNA binding.[7, 9, 10] They also showed that such an acidic extension improved the affinity of a peptide rationally designed to heterodimerize with human bZIP CEBPA.[8] Because the basic region of BZLF1 is highly similar to that of human bZIPs (Fig. 1a), we reasoned that incorporating an acidic extension into the N-terminus of our BD_{CC}^{245} design might enhance its affinity for BZFL1.

Three acidic extension variants developed by Vinson et al. differ in 2 positions that could interact with the BZLF1 basic region, if the interaction occurred with a coiled-coil-like geometry as has been hypothesized for other systems.[7] We chose to use the “A”-extension, which introduced the possibility of an attractive Glu-Arg *g*-*e′* interaction and a Leu-Leu

core-core *a-a′* interaction. Following prior work in the Vinson laboratory,[9] we constructed $A-BD_{CC}^{245}$ (sequence in Table 1). The modification added 17 residues at the N-terminus and replaced 6 out of 9 of the most N-terminal residues of the designed region (Table 1). Interestingly, $A-BD_{CC}^{245}$ showed much greater helicity than BD_{CC}^{245} and BD_{CC}^{231}, indicating that either some of the N-terminal 26 residues and/or the distal C-terminal region are likely helical in this context (Fig. 3f). The T_m for $A-BD_{CC}^{245}$ was similar to those for BD_{CC}^{231} and BD_{CC}^{245} (Table 1), whereas interaction with B-BZLF1^{245} was significantly stabilized compared to the BD_{CC}^{245}/$B-BZLF1^{245}$ interaction as expected (Fig. 3f). The heterocomplex melted at > 80 °C (Table 2). Together these observations indicate that changes made in $A-BD_{CC}^{245}$ did not stabilize the design homodimer, but further enhanced its interaction with B-BZLF1^{245}, as desired for inhibitor design.

For comparison, we constructed several other peptides with acidic extensions and assessed their self-association (Table 1). This modification dramatically destabilized BZLF1^{245} by 28 °C (71 °C for BZLF1^{245} vs. 43°C for A-BZLF1^{245}). A-BZLF1^{231} was also destabilized, but by only 10 °C (43°C for BZLF1^{231} vs. 33 °C for A-BZLF1^{231}). BD_{IED}^{245} was destabilized by an amount that could not be quantified because $A-BD_{IED}^{245}$ did not exhibit a cooperative melt. A-BZLF1^{245} was tested for interaction with B-BZLF1^{245} and formed a heterocomplex with T_m of 74 °C (compared to the T_m for B-BZLF1^{245} self-interaction, 67 °C, Tables 1, 2). The T_m for the heterocomplex between A-BZLF1^{231} and B-BZLF1^{231} was 58 °C (compared to the T_m for B-BZLF1^{231} self-interaction, 31 °C). These results are consistent with an interaction between the acidic extension and the basic region stabilizing the heterocomplexes, and also with an unfavorable interaction between the distal CT and the acidic extension, which is considered further in the Discussion.

Inhibiting DNA binding by BZLF1

We used an electrophoretic mobility shift assay (EMSA) to assess inhibition of B-BZLF1^{245} binding to DNA by different designed peptides (Fig. 6). The dimerization domain of BZLF1 lacking the basic region, BZLF1^{245}, was included for comparison purposes. All peptides tested showed concentration-dependent inhibition. BD_{CC}^{245}, A-BZLF1^{245} and $A-BD_{CC}^{245}$ were more effective than BZLF1^{245}. Design BD_{IED}^{245} was also an effective inhibitor. The most potent inhibitor was BD_{CC}^{245}, which completely inhibited B-BZLF1^{245} binding to DNA at equi-molar concentration.

Discussion

In this study, we employed different design strategies to create inhibitor peptides targeting the viral bZIP protein BZLF1. We sought peptides that achieved hetero-specificity through enhanced affinity for the target and/or reduced self-interaction. Below we discuss our different design approaches and the experimental behaviors of our designed peptides.

Applying CLASSY to BZLF1

As demonstrated earlier,[18] CLASSY is an algorithm that can be applied to design bZIP-like coiled coils. It was developed in conjunction with a specialized scoring function that includes computed structure-based terms, helix propensities, and experimentally determined coupling energies. The scoring function was validated on a large-scale dataset of human bZIP coiled-coil interactions[13] and supported the successful design of numerous bZIP-binding peptides. It is not known to what extent the bZIP scoring function can be applied in

design problems involving coiled-coil targets with features not observed in typical human bZIPs. Here, we explored whether the BZLF1 dimerization domain could be treated as a standard bZIP target for CLASSY design.

To treat BZLF1 as a coiled coil, we designed against the N-terminal part of the sequence and did experimental tests using constructs that did not include the distal CT (the "231" constructs, Fig 1b, 2a), much of which is not observed in the X-ray structure. The BZLF1 coiled-coil region is rather short (4 heptads), has only one Leu at position *d* among these heptads, and includes a region with very narrow inter-helical distance (~4 Å Cα-Cα distance at *a*-position residue 204). These variations might be expected to compromise performance of the scoring function, as coiled-coil context is known to influence the contributions of residues and residue pairs to stability.[17, 37, 38] Thus, methods validated using human bZIPs might not generalize broadly to all coiled-coil dimers. However, we found that design BD^{231}_{CC} incorporated elements very commonly employed in published anti-human bZIP designs (see below), and that these gave good experimental performance in this less canonical example. Success might be attributed to the fact that introducing more canonical residues at interfacial sites on one helix (the design) makes the design-target heterodimer more similar to the human bZIPs, e.g. the heterodimer likely has a more typical helix-helix separation.

Features contributing to the stability and specificity of the designs

Analysis of the designed sequences suggests that stability and specificity were achieved using different combinations of core, edge and core-edge interactions. For example, in the BD^{231}_{CC} design, the *a* and *d* heptad positions were populated with hydrophobic Ile and Leu, respectively, (e.g. Y200L, A204I, K207L), which are expected to be exceptionally stabilizing in the design homodimer.[39] Therefore, a strategy that used only these mutations to stabilize the design-target interaction would likely stabilize the design self-interaction even more, and fail to achieve heterospecificity. Negative design elements that likely compensate for over-stabilization of the design self-interaction come from interfacial *e* and *g* positions occupied by negatively charged amino acids. These negative charges make favorable interactions with positively charged residues in the target (e.g. 201R, 207K), consistent with improving the stability of the design-target interaction. However, they also introduce repulsive *g*-*e*′ or *e*–*g*′ interactions in the design homodimer (e.g. 196E–201E (*g*-*e*′), 203E–208E (*g*-*e*′), 201E–203E (*e*–*g*′)). Similar examples of using a highly hydrophobic core to achieve stability while modulating specificity using interfacial charge have been observed in many prior coiled-coil designs.[32] One less familiar feature in the BD^{231}_{CC} design is the presence of an N-terminal glutamate at a *d* position. Two glutamate residues at *d* and *d*′ in a homodimer are destabilizing in coiled coils,[40] but this residue potentially interacts favorably with an *e*′ lysine in BZLF1, via a core-to-edge type interaction that has previously been noted in CLASSY-derived designs and other studies.[17, 18, 41, 42, 43]

Designs BD^{231}_{IED} and BD^{231}_{EEI} relied much more on core-to-edge interactions, which were placed close to the middle of the coiled coil in these designs. In contrast to *g*-*e*′ interactions, no coupling energies have been measured for negatively charged residues at *d*-*d*′ or *d*–*e*′ sites. CLASSY performed poorly in predicting the relative stabilities of complexes involving BD^{231}_{IED} and BD^{231}_{EEI}, most likely because experimental data describing such charged core-core and core-edge interactions were not available to guide the development of the scoring function.[13,18] Nevertheless, a cluster of 4 negatively charged residues in the design homodimer proved very effective as a negative design element; BD^{231}_{IED} and BD^{231}_{EEI} did not appreciably self-associate. Affinity for the target was also compromised, however. Substitution of alanine with isoleucine at *a* position 204 was introduced to compensate for

some of the lost stability of the heterodimer, showing how a different combination of stabilizing and destabilizing elements can generate a hetero-specific design that inhibits DNA binding (Fig. 6).

Substitution of isoleucine for alanine at *a* position 204 is found in all 3 designs. In the native structure, alanine at this position fits well in the tight space between unusually close helices (~4 Å Cα-Cα distance between residue 204 on the two chains). Isoleucine cannot be built into this site in the crystal structure without severe clashes. Nonetheless, the larger Ile was accommodated in all three designs, and an alanine to isoleucine mutation is stabilizing in the context of BZLF1^{245} (an increase of T_m by 9 °C under the conditions of Table 1, data not shown). These data suggest a change in the backbone structure upon making this substitution. Local rearrangement of the design-BZLF1 complex to a more typical backbone structure probably helps explain why the CLASSY bZIP scoring function worked well. To achieve good predictive ability for a wider range of backbone structures, backbone flexibility could be treated explicitly.[43, 44]

The influence of the distal CT region

Previous studies revealed that the distal CT, although unresolved in the BZLF1 crystal structure, might interact with the N-terminal part of the BZLF1 coiled-coil region, thereby stabilizing the dimer.[27] We confirmed a stabilizing role for this region (Table 1, comparing BZLF1^{231} and BZLF1^{245}). Interestingly, this effect depends on the sequence in the coiled-coil region (Table 1, 2). The distal CT does not stabilize the BD_{CC}^{245} design self-interaction, and it enhances the stability of the BD_{CC}^{245}-target interaction only modestly. On the other hand, the distal CT significantly increased the stability of the BD_{IED}^{245} design self-interaction, as well as the stability of the BD_{IED}^{245}-target interaction. There are more sequence changes in the BD_{CC} design, and the number of negative charges introduced is larger than in the BD_{IED}, design. As discussed below, the influence of the distal CT is also sensitive to the acidic extension included in some designs. Although the structure of the interaction between the distal CT and the N-terminal part of the coiled coil in the native protein is not known, repulsive electrostatics, or unfavorable desolvation of charges in the coiled-coil region are plausible mechanisms for disfavoring this interaction in the BD_{CC} design.

Specificity against human bZIPs

We did not consider specificity against human bZIPs in our design procedure. However, we showed that the design BD_{CC} is not promiscuous in binding human bZIP proteins. Computational analysis predicted that the coiled-coil region of BD_{CC} would interact with the BZLF1 coiled coil moderately more favorably than with any other human bZIP coiled coil (but with a few close competitors). This is interesting, given the fairly canonical coiled-coil sequence features of BD_{CC}. The requirement to satisfy hydrogen bonding for Asn 204 at the *a* position in BD_{CC}, and the charge complementarity between the *e* and *g* positions of BD_{CC} and BZLF1 helices but not most human proteins, contributed to the predicted binding preference.

Thermal stability studies confirmed that BD_{CC}^{245} does not bind strongly to selected human bZIPs identified in the computational analysis. In addition to selectivity derived from the coiled-coil region (which was predicted to be modest), the CT region likely confers additional specificity. Interactions with BD_{CC}^{245} and B-BZLF1^{245} could benefit from native-like contacts between the CT region and the coiled coil domain, which are not conserved in

complexes with human proteins. Thus, the interaction specificity of BD_{CC}^{245} is likely encoded in both its coiled-coil domain and the CT region.

Improving inhibitor potency using an N-terminal acidic extension

The Vinson group has demonstrated that dominant-negative inhibitors of bZIP dimerization and DNA binding can be created by replacing the basic region of native or modified native bZIPs with an acidic sequence.[7] In this study, we used this strategy to improve the potency of our designed peptides. The resulting $A-BD_{CC}^{245}$ peptide maintained specificity, showing little change in the T_m for the design self-association. The small change in homodimer stability probably results from destabilization by the negative charges in the extension, countered by a stabilizing leucine residue introduced at *d* position 193 (this residue is Glu in BD_{CC}).[9] $A-BD_{CC}^{245}$ formed a more stable complex with the target B-BZLF1^{245} than did BD_{CC}^{245} (an increase of T_m > 14 °C at 4 μM, Table 2). This indicates that the acidic extension, which targets the basic region of bZIPs, can be used in conjunction with computational design methods targeting the coiled coil. Given that the Vinson laboratory has demonstrated that the coiled-coil region of A-ZIPs governs interaction specificity, while the acidic extension provides much enhanced affinity, this is an appealing strategy for expanding the design of tight-binding and selective bZIP inhibitors.[7, 8, 9, 10, 18]

Interestingly, modifying BZLF1 with an acidic extension did not stabilize interaction of A-BZLF1^{245} with B-BZLF1^{245} as much as expected (T_m of 74 °C compared to 67 °C for the B-BZLF1^{245} homodimer, Table 1, 2). In contrast, interaction of the shorter construct A-BZLF1^{231} with B-BZLF1^{231} was stabilized to a much greater extent (T_m of 58 °C compared to 31 °C for the B-BZLF1^{231} homodimer). Furthermore, the destabilizing effect of the acidic extension on design homodimer stability is quite different in BZLF1^{245} vs. BZLF1^{231} (decreasing T_m values by 28 °C vs. 10 °C, Table 1). These observations are consistent with a model where the distal CT interacts unfavorably with the acid extension, much as it appears to interact unfavorably with negative charges in the N-terminal part of the BD_{CC} design. Although not addressed in the present study, the performance of A-BZLF1^{245} as an inhibitor could potentially be improved by redesigning the acidic extension so that interference from the distal CT is minimized, although this is difficult in the absence of structural information about this part of the protein.

Analysis of inhibitor potency

To test the designed peptides as inhibitors of BZLF1 DNA binding, we used an *in vitro* EMSA assay to monitor the population of B-BZLF1^{245} bound to DNA in the presence of different peptides (Fig. 6). It is unsurprising that $A-BD_{CC}^{245}$, which formed the most thermostable complex with B-BZLF1^{245} and exhibited the largest difference in homodimer vs. heterodimer stability, was the most potent inhibitor. The improved performance of BD_{CC}^{245} and A-BZLF1^{245} relative to the native peptide, BZLF1^{245}, could be rationalized by their improved affinity and/or anti-homodimer specificity (see below). BD_{IED}^{245} inhibited DNA binding effectively and we estimate its potency is similar to that of BZLF 1^{245}, although these two peptides could not be compared using identical assay conditions (see Materials and Methods). The effectiveness of BD_{IED}^{245} resulted from a combination of reduced affinity but improved anti-homodimer specificity.

To explore more generally how affinity and specificity each influence potency, we constructed a simple computational model with the following assumptions: 1) the target bZIP, the DNA, and the designed peptide were the only components present, 2) the target

bZIP homodimer was the only species that could bind DNA (i.e. complete cooperative binding), 3) non-specific DNA binding was neglected. Some of the assumptions made may not apply to all of our experiments. We computed concentration dependent inhibition of DNA binding for a series of designs covering a spectrum of affinities and specificities. Affinity was described by the ratio between the dissociation constant of the target bZIP homodimer and that of the design-target heterodimer (K_d^{T2}/K_d^{DT}, D: design, T: target, see Materials and Methods), and specificity was described by the ratio between the dissociation constant for the design homodimer and that of the design-target heterodimer (K_d^{D2}/K_d^{DT}). The efficacy of different inhibitors is illustrated in a heat map in Fig. 7 that indicates the improvement in IC50 over a reference for which $K_d^{D2} = K_d^{DT} = K_d^{T2}$. The reference inhibitor with affinity and specificity of 1 was included to reflect the behavior of the dimerization domain of the target bZIP. We explored two scenarios that led to different inhibition landscapes: one where modeled dissociation constants for the target bZIP complex and bZIP-DNA interactions were lower than the target bZIP concentration (Fig. 7a), and another where they were higher (Fig.7b)

The results in Fig. 7 support intuition about the importance of both affinity and specificity. Lines of constant color running across the plots in Fig. 7 show that equivalent potency can be achieved using different combinations of affinity and specificity. Clearly, neither affinity nor preference for hetero vs. homodimerization correlates directly with design performance. For the purposes of discussion, we label 3 regions on the plots: $H_{affinity}$:L_{spec} indicates inhibitors with high affinity for the target but limited anti-homodimer specificity, $L_{affinity}$:H_{spec} indicates inhibitors with affinity for the target that is comparable to or weaker than the reference inhibitor, but with weaker self-association, and $H_{affinity}$:H_{spec} inhibitors have both tighter target-binding affinity and weaker self-association than the reference. Among our designs, and to the extent that approximate stabilities assessed by thermal denaturation under CD conditions can be extrapolated to the gel-shift assay, BD_{CC}^{245} and BD_{IFD}^{245} are both $L_{affinity}$:H_{spec} inhibitors that use anti-homodimer specificity to improve inhibitor potency. A – BD_{CC}^{245} maintains anti-homodimer specificity but gains additional affinity via the acidic extension, making it a $H_{affinity}$:H_{spec} inhibitor.

The model in Fig. 7 is useful for broadly guiding the computational design of specific inhibitors, so we conclude with a few general observations. First, heterospecificity is important, but not sufficient, for good performance. A design is hetero-specific if the ratio $K_d^{T2} \bullet K_d^{D2}/(K_d^{DT})^2$ is larger than 1. In the figure, this region is below the dashed line and all inhibitors with potency better than the reference lie in this region. Maintaining hetero-specificity for high affinity designs imposes a bound on design homodimer stability. This is relevant for parallel dimeric coiled-coil targets, because amino-acid changes that enhance interaction with the target often stabilize the design self-interaction even more.[39] Second, the relative importance of improving affinity vs. specificity depends on the target and assay conditions. For panel a, improved hetero-specificity implies enhanced design performance regardless of whether affinity or specificity is the main contributor. On the other hand, if the target bZIP concentration is lower, as in panel b, improving specificity alone is no longer sufficient, and affinity must be optimized; very potent designs in panel b can only be achieved by optimizing along the path toward $H_{affinity}H_{spec}$. Finally, the overall diagonal trends for constant-IC_{50} regions in both panels emphasize that improving either affinity or specificity can potentially lead to success, depending on the specific conditions and requirements for an application. Designs belonging to the $H_{affinity}$ H_{spec} class are the most effective. However, such designs might not exist, or could be hard to identify for a particular problem. In such cases, one could consider optimizing primarily affinity or specificity, depending on which is easier to achieve. Although not used extensively for this purpose

here, the CLASSY algorithm is well suited for identifying designs with different affinity vs. specificity trade-offs.[18]

Conclusion: implications for protein design

This study addresses three topics relevant to the design of peptides that inhibit native protein-protein interactions. First is the issue of specificity, which arises in many protein design problems and is acute for coiled-coil targets where self-association of the design can compete with target inhibition. Using BZLF1 as a target, we characterized peptides that balance affinity and specificity in different ways. This adds to the small number of examples where affinity and specificity have both been treated as design considerations.[18, 41, 43, 45, 46, 47, 48, 49] Second, we explored a design problem where features of the target that are not well described in an existing structure (the BZLF1 distal CT) nevertheless influence complex stability. We showed that different designs responded differently to the introduction of the distal CT. This argues for developing methods that broadly survey design solution space and discovering a large set of potentially good designs, rather than identifying only “the best” design according to some imperfect criteria. This can be accomplished in various ways, e.g. by exploring a range of tradeoffs between stability and specificity, or exploring a variety of related structural templates as design scaffolds.[18, 50] Testing diverse solutions maximizes the chance of finding a design that interacts well with poorly characterized features of the target. Finally, our best design exploited a modular strategy where optimization of the coiled-coil dimerization interface was coupled with a more generic strategy developed previously for stabilizing inhibitor-bZIP complexes. Modularity is likely to be a key strategy for the design of ever more complex molecular parts.

Materials and Methods

Cloning, protein expression and purification

Synthetic genes encoding native or redesigned BZLF1 sequence, residues 175 or 191 to 245 (B-BZLF1^{245}, BZLF1^{245}, BD^{231}_{CC}, BD^{245}_{IED}), were constructed by gene synthesis. Primers were designed using DNAWorks,[51] and a two-step PCR procedure was used for annealing and amplification. Genes encoding the native or redesigned sequence in the context of residues 191 to 231 were made in a single-step PCR reaction using the longer constructs as templates. The genes were cloned via BamHI/XhoI restriction sites into a modified version of a pDEST17 vector that encodes an N-terminal 6xHis tag and a GESKEYKKGSGS linker that improves the solubility of the recombinant protein.[28] To facilitate cloning of genes encoding the acidic extension, a pET16b vector (Novagen) was modified to encode an N-terminal 6xHis tag, followed by a GSY linker and the acidic extension sequence. Genes encoding BZLF1^{231}, BZLF1^{245} and the designs BD^{245}_{CC} and BD^{245}_{IED} were subsequently cloned into the modified vector using AflII/XhoI restriction sites to make A-BZLF1^{231}, A-BZLF1^{245}, A – BD^{245}_{CC} and A – BD^{245}_{IED}. Recombinant proteins were expressed in *E. coli* RP3098 cells. Cultures were grown at 37 °C to an OD of ~0.4–0.9, and expression was induced by addition of 1 mM IPTG. Purification was performed under denaturing conditions (6M GdnHCl) using an Ni-NTA affinity column followed by reverse-phase HPLC. Human bZIP constructs containing the basic region and the coiled-coil domain were described previously.[28]

Computational protein design using CLASSY

The sequence BD_{CC} was designed using the CLASSY algorithm as previously reported.[18] In brief, the algorithm solves for the sequence predicted to interact most favorably with a target sequence (here, chosen to be the N-terminal part of the BZLF1 leucine zipper, residues 191

to 209) using integer linear programming. It is possible to impose constraints on the gap between the energy of interaction with the target and the energy of undesired states such as the design homodimer. No such constraint was applied in the design of BD_{CC}, which was predicted to favor the design-target interaction over design homodimerization without it. The scoring function used was HP/S/Cv. This function was derived by combining molecular mechanics calculations and experimentally determined coupling energies for many core *a*-*a*′ interactions.[13, 16] The Leu-Leu core *d*-*d*′ interaction was modeled with an empirical value of −2 kcal/mol^{-1}. The HP/S/Cv structure-based energy function was transformed into a sequence-based expression using cluster expansion, and modified using empirical data, as described by Grigoryan et al.[18]

Predicting interactions between BD_{CC} and human bZIPs

BZLF1 was aligned with 36 human bZIPs using the conserved basic region, and interaction scores for residues 191–221 of BD_{CC} with the correspondingly aligned 31 residues of each human bZIP were computed using the HP/S/Cv model as described above.

Circular dichroism spectroscopy

Circular dichroism experiments were performed and analyzed, and T_m values fitted as described previously.[18] Thermal melts from 0 °C to 85 °C were mostly reversible, regaining ≥95% of signal or giving closely similar T_m values for the reverse melt (except for samples containing NFIL3, which precipitated upon heating to 85 °C). Melting temperatures were estimated by fitting the data to a two-state equilibrium (unfolded/folded), assuming no heat capacity changes upon folding. A detailed description of the equation was described previously. [18] In cases where high-temperature unfolding precluded accurate fitting of unfolded baselines, the T_m was either defined as the mid-point of the unfolding transition after manually picking the baseline (for the 1:1 mixture of B-$BZLF1^{245}$ and A-$BZLF1^{245}$), or a lower bound on the Tm value was estimated (for the 1:1 mixture of B-BZLF1 and A – BD_{CC}^{245}). The protein concentrations are given in the figure legends. All measurements were performed in PBS buffer containing 12.5 mM potassium phosphate (pH 7.4), 150 mM KCl, 0.25 mM EDTA and 1 mM DTT. Samples were heated to 65 °C for 5 minutes before measurement to equilibrate peptide mixtures, and then cooled to and equilibrated at the starting temperature.

Analytical ultracentrifugation

Protein samples were dialyzed against the reference buffer (12.5 mM sodium phosphate, 150 mM NaCl, 1mM DTT, 0.25 mM EDTA, pH 7.4) three times (including once overnight) before measurements. Sedimentation equilibrium runs were performed with a Beckman XL-I analytical ultracentrifuge using interference optics. Two concentrations for each protein sample were prepared (50 and 100 μM), and runs at 3 different speeds (28,000, 35,000 and 48,000 rpm) were carried out at 20 °C. Each run was ~ 20 h, and equilibrium was confirmed by negligible differences between the sample distribution in the cell over sequential scans. Data were analyzed globally with the program HeteroAnalysis[52], using a calculated[53] partial specific volume of 0.7275 ml/g (for the BD_{CC}^{231}/$BZLF1^{231}$ mixture) or 0.7245 ml/g (for BD_{CC}^{231}) and a solution density of 1.005 g/ml.

Electrophoretic mobility shift assay (EMSA)

Gel shift assays were performed as described previously[28]. Briefly, 10 nM B-$BZLF1^{245}$ was prepared either alone or mixed with each inhibitor at 9 concentrations ranging from 10 nM to 2560 nM in 2-fold dilutions. Gel-shift buffer ((150 mM KCl, 25 mM TRIS pH 8.0, 0.5 mM EDTA, 2.5 mM DTT, 1 mg/ml BSA, 10% (v/v) glycerol, 0.1 μg/ml competitor DNA

(Poly (I)·Poly (C) (Sigma))) was then added and incubated for 10 minutes at 42 °C. Closely similar results were obtained when incubating samples for 20 minutes at 42 °C. The competitor BD_{IED}^{245} was not stable upon heating and was incubated for 2 hours at 18–22 °C. Radiolabeled annealed AP-1 site, CGCTTGATGACTCAGCCGGAA (IDT), at a final concentration of 0.7 nM was added and incubated for 15 minutes at 18–22 °C. Complexes were separated on NOVEX DNA retardation gels (Invitrogen). Dried gels were imaged using a phosphorimaging screen and a Typhoon 9400 imager. ImageQuant software (Amersham Biosciences) was used to quantify band intensities.

Simulating the impact of affinity and specificity on designed peptide behaviors

The simulation treated the following species: The target bZIP monomer (T), the target bZIP homodimer (T_2), the design monomer (D), the design homodimer (D_2), the design-target bZIP heterodimer (DT), free DNA (DNA) and the complex formed between the target bZIP homodimer and DNA (T_2DNA). Species are linked by the following reactions:

$$2T \leftrightarrows T_2 \quad \frac{[T]^2}{[T_2]} = K_d^{T_2}$$
$$2D \leftrightarrows D_2 \quad \frac{[D]^2}{[D_2]} = K_d^{D_2}$$
$$D+T \leftrightarrows DT \quad \frac{[D][T]}{[DT]} = K_d^{DT}$$
$$T_2+DNA \leftrightarrows T_2DNA \quad \frac{[T_2][DNA]}{[T_2DNA]} = K_c^{T_2DNA}$$
$$[T]+[DT]+2[T_2]+2[T_2DNA]=[T]_{total}$$
$$[D]+[DT]+2[D_2]=[D]_{total}$$
$$[DNA]+[T_2DNA]=[DNA]_{total}$$

Affinity is defined as K_d^{T2}/K_d^{DT}, and a value > 1 indicates the design-target bZIP heterodimer is more stable than the target bZIP homodimer (improved affinity). Specificity is defined as K_d^{D2}/K_d^{DT}, and a value > 1 indicates the design-target bZIP heterodimer is more stable than design homodimer (improved specificity). A design with affinity and specificity equal to 1 was used as a reference. The IC_{50} value was defined as the design concentration $[D]_{total}$ at which 50% less DNA is bound relative to zero design concentration. The total target bZIP concentration $[T]_{total}$ was fixed at 10 nM, and the total DNA concentration $[DNA]_{total}$ at 0.7 nM. Different combinations of K_d^{T2} and K_d^{T2DNA} values were explored (10^{-9}, 10^{-8}, and 10^{-7} M for each), including when both are lower than $[T]_{total}$ (10^{-9} M/10^{-9} M, Fig. 7a) and when both are higher than $[T]_{total}$ (10^{-7} M/10^{-7} M, Fig. 7b). For each combination of fixed K_d^{T2} and K_d^{T2DNA}, the IC_{50} values for a range of designs with different affinities (0.1 to 10) and specificities (0.1 to 100) were calculated. The ratio $IC_{50}^{design}/IC_{50}^{ref}$, with a value < 1 implying greater potency than the reference, was plotted as a heat map. The dashed lines on the plots in Fig. 7 indicate points where the product of affinity and specificity ($(K_d^{T2} * K_d^{D2})/(K_d^{DT} * K_d^{DT})$) equals 1. All designs below the dashed line are hetero-specific. The simulation was carried out and heat maps were generated using Matlab (MathWorks).

Acknowledgments

We thank K. E. Thompson for designing the acidic extension vector, making the A-BZLF1^{245} construct, and providing valuable suggestions. We thank G. Grigoryan for assistance with the CLASSY algorithm, and members of the Keating lab, especially O. Ashenberg, C. Negron, S. Dutta, L. Reich, V. Potapov, K. Hauschild and J. DeBartolo for helpful discussion of the manuscript. We thank D. Pheasant at the Biophysical Instrumentation Facility at MIT for assistance in analytical ultracentrifugation experiments. A.W.R. was supported by a Koch graduate fellowship. This work was funded by NIH award GM067681 and used computer resources provided by NSF award 0821391.

Abbreviations used

bZIP	basic-region leucine zipper
BSA	bovine serum albumin
CD	circular dichroism
CC	coiled coil
CLASSY	cluster expansion and integer linear programming based analysis of specificity and stability
CT	C-terminal
DTT	dithiothreitol
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
GdnHCl	guanidine hydrochloride
HPLC	high-performance liquid chromatography
IPTG	isopropyl β-d-1-thiogalactopyranoside
PBS	phosphate-buffered saline
Ni-NTA	nickel nitrilotriacetic acid
T_m	melting temperature
TRIS	tris(hydroxymethyl)aminomethane

References

1. O'Shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science. 1991; 254:539–544. [PubMed: 1948029]
2. Rishi V, Potter T, Laudeman J, Reinhart R, Silvers T, Selby M, Stevenson T, Krosky P, Stephen AG, Acharya A, Moll J, Oh WJ, Scudiero D, Shoemaker RH, Vinson C. A high-throughput fluorescence-anisotropy screen that identifies small molecule inhibitors of the DNA binding of B-ZIP transcription factors. Anal Biochem. 2005; 340:259–271. [PubMed: 15840499]
3. Rishi V, Oh WJ, Heyerdahl SL, Zhao J, Scudiero D, Shoemaker RH, Vinson C. 12 Arylstibonic acids that inhibit the DNA binding of five B-ZIP dimers. J Struct Biol. 2010; 170:216–225. [PubMed: 20176111]
4. Mason JM, Schmitz MA, Muller KM, Arndt KM. Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design. Proc Natl Acad Sci U S A. 2006; 103:8989–8994. [PubMed: 16754880]
5. Mason JM, Muller KM, Arndt KM. Positive aspects of negative design: simultaneous selection of specificity and interaction stability. Biochemistry. 2007; 46:4804–4814. [PubMed: 17402748]
6. Mason JM, Hagemann UB, Arndt KM. Role of hydrophobic and electrostatic interactions in coiled coil stability and specificity. Biochemistry. 2009; 48:10380–10388. [PubMed: 19743874]
7. Acharya A, Rishi V, Moll J, Vinson C. Experimental identification of homodimerizing B-ZIP families in Homo sapiens. J Struct Biol. 2006; 155:130–139. [PubMed: 16725346]
8. Krylov D, Olive M, Vinson C. Extending dimerization interfaces: the bZIP basic region can form a coiled coil. EMBO J. 1995; 14:5329–5337. [PubMed: 7489722]
9. Olive M, Krylov D, Echlin DR, Gardner K, Taparowsky E, Vinson C. A dominant negative to activation protein-1 (AP1) that abolishes DNA binding and inhibits oncogenesis. J Biol Chem. 1997; 272:18586–18594. [PubMed: 9228025]

10. Ahn S, Olive M, Aggarwal S, Krylov D, Ginty DD, Vinson C. A dominant-negative inhibitor of CREB reveals that it is a general mediator of stimulus-dependent transcription of c-fos. Mol Cell Biol. 1998; 18:967–977. [PubMed: 9447994]
11. Gerdes MJ, Myakishev M, Frost NA, Rishi V, Moitra J, Acharya A, Levy MR, Park SW, Glick A, Yuspa SH, Vinson C. Activator protein-1 activity regulates epithelial tumor cell identity. Cancer Res. 2006; 66:7578–7588. [PubMed: 16885357]
12. Oh WJ, Rishi V, Orosz A, Gerdes MJ, Vinson C. Inhibition of CCAAT/enhancer binding protein family DNA binding in mouse epidermis prevents and regresses papillomas. Cancer Res. 2007; 67:1867–1876. [PubMed: 17308129]
13. Grigoryan G, Keating AE. Structure-based prediction of bZIP partnering specificity. J Mol Biol. 2006; 355:1125–1142. [PubMed: 16359704]
14. Fong JH, Keating AE, Singh M. Predicting specificity in bZIP coiled-coil protein interactions. Genome Biol. 2004; 5:R11. [PubMed: 14759261]
15. Krylov D, Mikhailenko I, Vinson C. A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. EMBO J. 1994; 13:2849–2861. [PubMed: 8026470]
16. Acharya A, Rishi V, Vinson C. 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. 2006; 45:11324–11332. [PubMed: 16981692]
17. Steinkruger JD, Woolfson DN, Gellman SH. Side-chain pairing preferences in the parallel coiled-coil dimer motif: insight on ion pairing between core and flanking sites. J Am Chem Soc. 2010; 132:7586–7588. [PubMed: 20465308]
18. Grigoryan G, Reinke AW, Keating AE. Design of protein-interaction specificity gives selective bZIP-binding peptides. Nature. 2009; 458:859–864. [PubMed: 19370028]
19. Newman JR, Keating AE. Comprehensive identification of human bZIP interactions with coiled-coil arrays. Science. 2003; 300:2097–2101. [PubMed: 12805554]
20. Vinson C, Acharya A, Taparowsky EJ. Deciphering B-ZIP transcription factor interactions in vitro and in vivo. Biochim Biophys Acta. 2006; 1759:4–12. [PubMed: 16580748]
21. Countryman J, Jenson H, Seibl R, Wolf H, Miller G. Polymorphic proteins encoded within BZLF1 of defective and standard Epstein-Barr viruses disrupt latency. J Virol. 1987; 61:3672–3679. [PubMed: 2824806]
22. Schepers A, Pich D, Hammerschmidt W. A transcription factor with homology to the AP-1 family links RNA transcription and DNA replication in the lytic cycle of Epstein-Barr virus. EMBO J. 1993; 12:3921–3929. [PubMed: 8404860]
23. Feederle R, Kost M, Baumann M, Janz A, Drouet E, Hammerschmidt W, Delecluse HJ. The Epstein-Barr virus lytic program is controlled by the cooperative functions of two transactivators. EMBO J. 2000; 19:3080–3089. [PubMed: 10856251]
24. Liu P, Speck SH. Synergistic autoactivation of the Epstein-Barr virus immediate-early BRLF1 promoter by Rta and Zta. Virulogy. 2003; 310:199–206.
25. Young LS, Rickinson AB. Epstein-Barr virus: 40 years on. Nat Rev Cancer. 2004; 4:757–768. [PubMed: 15510157]
26. Petosa C, Morand P, Baudin F, Moulin M, Artero JB, Muller CW. Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein. Mol Cell. 2006; 21:565–572. [PubMed: 16483937]
27. Schelcher C, Al Mehairi S, Verrall E, Hope Q, Flower KBB, Woolfson DN, West MJ, Sinclair AJ. Atypical bZIP domain of viral transcription factor contributes to stability of dimer formation and transcriptional function. J Virol. 2007; 81:7149–7155. [PubMed: 17459922]
28. Reinke AW, Grigoryan G, Keating AE. Identification of bZIP interaction partners of viral proteins HBZ, MEQ, BZLF1, and K-bZIP using coiled-coil arrays. Biochemistry. 2010; 49:1985–1997. [PubMed: 20102225]
29. Hicks MR, Al-Mehairi SS, Sinclair AJ. The zipper region of Epstein-Barr virus bZIP transcription factor Zta is necessary but not sufficient to direct DNA binding. J Virol. 2003; 77:8173–8177. [PubMed: 12829857]

30. O'Shea EK, Lumb KJ, SKP. Peptide 'Velcro': design of a heterodimeric coiled coil. Curr Biol. 1993; 3:658–667. [PubMed: 15335856]
31. Vinson C, Myakishev M, Acharya A, Mir AA, Moll JR, Bonovich M. Classification of human B-ZIP proteins based on dimerization properties. Mol Cell Biol. 2002; 22:6321–6335. [PubMed: 12192032]
32. Woolfson DN. The design of coiled-coil structures and assemblies. Adv Protein Chem. 2005; 70:79–112. [PubMed: 15837514]
33. Grigoryan G, Keating A. Structural specificity in coiled-coil interactions. Curr Opin Struct Biol. 2008; 18:477–483. [PubMed: 18555680]
34. Harbury PB, Zhang T, Kim PS, Alber T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science. 1993; 262:1401–1407. [PubMed: 8248779]
35. Taylor CM, Keating AE. Orientation and oligomerization specificity of the Bcr coiled-coil oligomerization domain. Biochemistry. 2005; 44:16246–16256. [PubMed: 16331985]
36. Mason JM, Arndt KM. Coiled coil domains: stability, specificity, and biological implications. Chembiochem. 2004; 5:170–176. [PubMed: 14760737]
37. Moitra J, Szila'k L, Krylov D, Vinson C. Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. Biochemistry. 1997; 36:12567–12573. [PubMed: 9376362]
38. Lu SM, Hodges RS. Defining the minimum size of a hydrophobic cluster in two-stranded alpha-helical coiled-coils: effects on protein stability. Protein Sci. 2004; 13:714–726. [PubMed: 14978309]
39. Acharya A, Ruvinov SB, Gal J, Moll JR, Vinson C. 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. 2002; 41:14122–14131. [PubMed: 12450375]
40. Tripet B, Wagschal K, Lavigne P, Mant CT, Hodges RS. 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. 2000; 300:377–402. [PubMed: 10873472]
41. Havranek JJ, Harbury PB. Automated design of specificity in molecular recognition. Nat Struct Biol. 2003; 10:45–52. [PubMed: 12459719]
42. Reinke AW, Grant RA, Keating AE. A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. J Am Chem Soc. 2010; 132:6025–6031. [PubMed: 20387835]
43. Barth P, Schoeffler A, Alber T. Targeting metastable coiled-coil domains by computational design. J Am Chem Soc. 2008; 130:12038–12044. [PubMed: 18698842]
44. Mandell DJ, Kortemme T. Backbone flexibility in computational protein design. Curr Opin Biotechnol. 2009; 20:420–428. [PubMed: 19709874]
45. Kortemme T, Joachimiak LA, Bullock AN, Schuler AD, Stoddard BL, Baker D. Computational redesign of protein-protein interaction specificity. Nat Struct Biol. 2004; 11:371–379.
46. Ali MH, Taylor CM, Grigoryan G, Allen KN, Imperiali B, Keating AE. Design of a heterospecific, tetrameric, 21-residue miniprotein with mixed alpha/beta structure. Structure. 2005; 13:225–234. [PubMed: 15698566]
47. Bolon DN, Grant RA, Baker TA, Sauer RT. Specificity versus stability in computational protein design. Proc Natl Acad Sci USA. 2005; 102:12724–12729. [PubMed: 16129838]
48. Sammond DW, Eletr ZM, Purbeck C, Kuhlman B. Computational design of second-site suppressor mutations at protein-protein interfaces. Proteins. 2010; 78:1055–1065. [PubMed: 19899154]
49. Karanicolas J, Kuhlman B. Computational design of affinity and specificity at protein-protein interfaces. Curr Opin Struct Biol. 2009; 19:458–463. [PubMed: 19646858]
50. Fu X, Apgar JR, Keating AE. Modeling backbone flexibility to achieve sequence diversity: the design of novel alpha-helical ligands for Bcl-xL. J Mol Biol. 2007; 371:1099–1117. [PubMed: 17597151]
51. Hoover DM, Lubkowski J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. Nucleic Acids Res. 2002; 30:e43. [PubMed: 12000848]
52. Cole, JL.; Lary, JW. Heteroanalysis, Analytical Ultracentrifugation Facility. University of Connecticut; Storrs, CT: 2006.

53. Laue, TM.; Shah, BD.; Ridgeway, TM.; Pelletier, SL. Analytical Ultracentrifugation in Biochemistry and Polymer Science. Cambridge: Royal Society of Chemistry; 1992. Computer aided interpretation of analytical sedimentation data for proteins; p. 90-125.

54. Glover JN, Harrison SC. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. Nature. 1995; 373:257–261. [PubMed: 7816143]

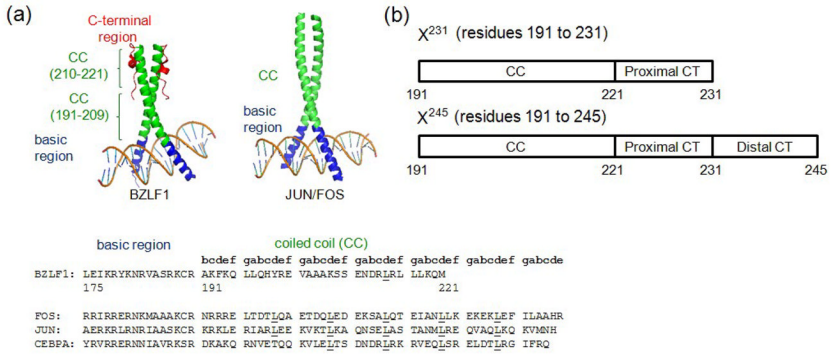

Figure 1.

Sequence and structure of the BZLF1 bZIP domain. (a) Crystal structure of BZLF1 bound to DNA[26] (PDB ID 2C9L, left) compared to human JUN/FOS bound to DNA[54] (PDB ID 1FOS, right). The basic region is blue, the coiled coil is green, and the C-terminal (CT) region is red. At the bottom are sequence alignments for the basic and coiled-coil regions of BZLF1 and representative human bZIPs. Leucines at *d* positions in the coiled coils are underlined. (b) Scheme of constructs used in this study. The “231” construct includes the coiled coil (CC) and the proximal C-terminal (CT) region; the “245” construct includes the coiled coil (CC) and the full-length C-terminal (CT) region.

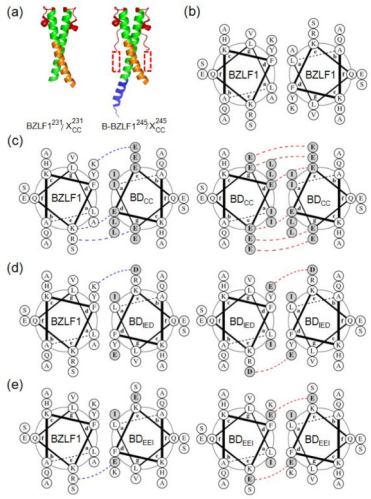

Figure 2.

Designed inhibitors. (a) Structural models representing two types of design-BZLF1 complexes tested in this work. At left, the “231” constructs, and at right, the “245” constructs. “X” is a placeholder for the name of a design, e.g. BD_{CC}. Color is as in Fig. 1a except that the designed region is shown in orange. The dashed boxes in the “245” complex indicate that part of the distal CT (237–245) is not resolved in the crystal structure. (b) Helical wheel diagram for the BZLF1 homodimer. (c–e) Helical wheel diagrams for the designs. On the left are design-target heterodimers and on the right are design homodimers. Design residues are highlighted in bold and with a grey background. Potential electrostatic interactions are indicated in blue if attractive and red if repulsive, (c) Design BD_{CC}, (d) Design BD_{IED}, (e) Design BD_{EEI}. In all helical wheel diagrams, only residues from *b* position 191 (Ala) to *f* position 209 (Ser) are shown (this region is orange in Fig. 2a), with the helix proceeding from N-to-C terminus into the page. Diagrams generated using DrawCoil 1.0 (http://www.gevorggrigoryan.com/drawcoil/).

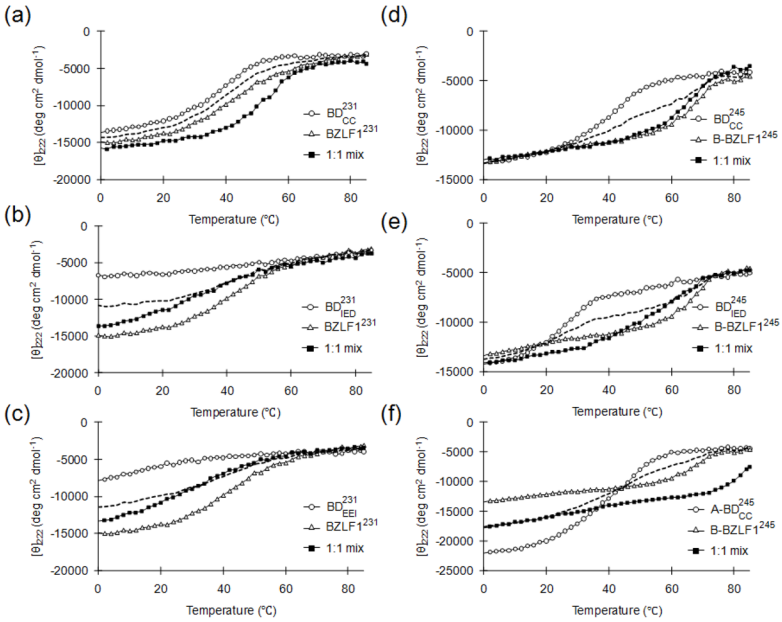

Figure 3.

Melting curves for targets, designs and complexes monitored by mean residue ellipticity at 222 nm. Four curves are shown in each panel: the target at 4 μM (open triangles), the design at 4 μM (open circles), a mixture of the target and the design at 2 μM each (closed squares), and the numerical average of the individual melting curves for the target and the design (short dashed lines). The target is $BZLF1^{231}$ for panels (a)–(c) and B-$BZLF1^{245}$ for panels (d)–(f), as described in text, and the designs are: (a) BD^{231}_{CC}, (b) BD^{231}_{IED}, (c) BD^{231}_{EEI}, (d) BD^{245}_{CC}, (e) BD^{245}_{IED}, and (f) A – BD^{245}_{CC}.

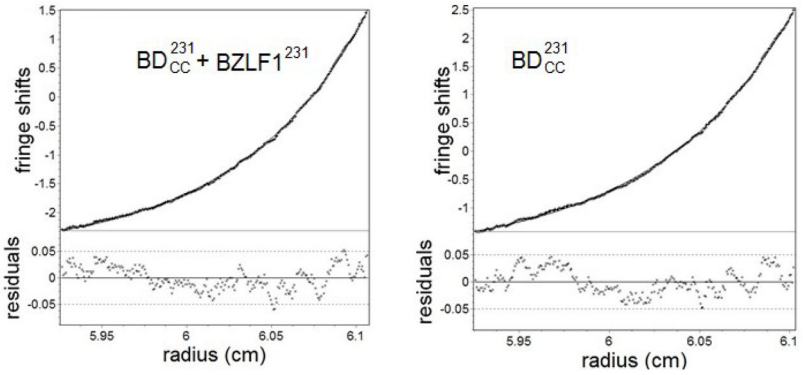

Figure 4.

Representative analytical ultracentrifugation data for BD_{CC}^{231}+$BZLF1^{231}$ (left) and BD_{CC}^{231} (right). The fits shown were obtained with data collected at 2 concentrations and 3 different centrifuge speeds. At the bottom are the residuals to the fit.

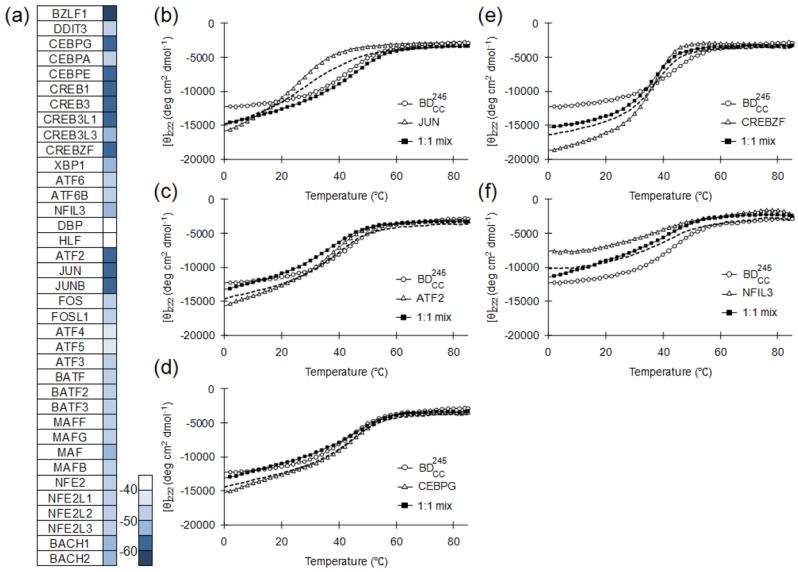

Figure 5.

(a) Predicted scores for BD_{CC} interacting with BZLF1 or human bZIP peptides, (b–f) Melting curves for selected human bZIP peptides, BD_{CC}^{245} or 1:1 mixtures of the two, monitored by mean residue ellipticity at 222 nm. Four curves are shown in each panel: the human bZIP at 4 μM (open triangles), BD_{CC}^{245} at 4 μM (open circles), a mixture of the human protein and BD_{CC}^{245} at 2 μM each (closed squares), and the numerical average of the individual melting curves for the human bZIP and the design (short dashed lines). The human bZIPs are: (b) JUN, (c) ATF2, (d) CEBPG, (e) CREBZF, and (f) NFIL3.

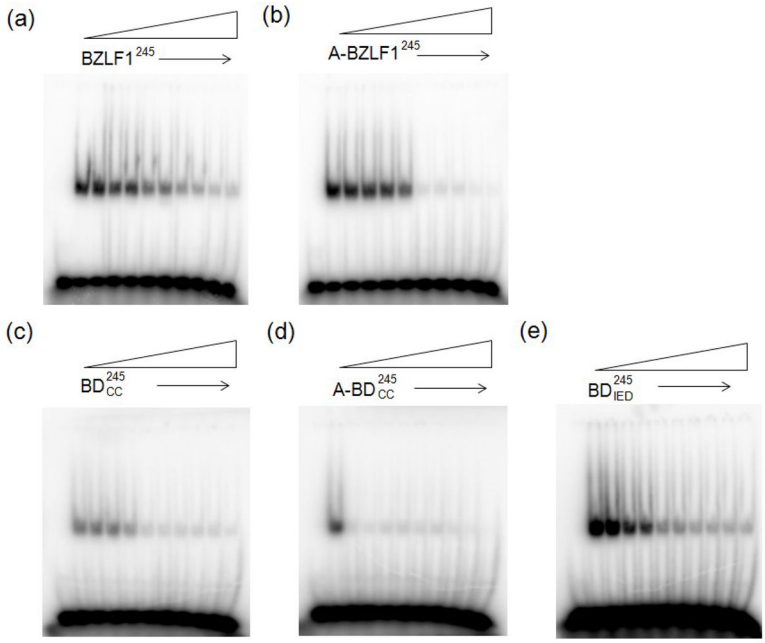

Figure 6.

Peptide inhibition of B-BZLF1^{245} binding to DNA. Representative gel-shift images were shown for: (a) BZLF1^{245}, (b) A-BZLF1^{245}, (c) BD_{CC}^{245}, (d) A – BD_{CC}^{245}, (e) BD_{IED}^{245}. The first two lanes for each gel include DNA only (first lane) and B-BZLF1^{245} with DNA (second lane). Inhibitor peptides were added in increasing concentrations from 10 nM to above 2 μM (left to right, 2-fold dilutions). Conditions are described in Materials and Methods in more detail, and were slightly different for panel (a)–(d) vs. panel (e).

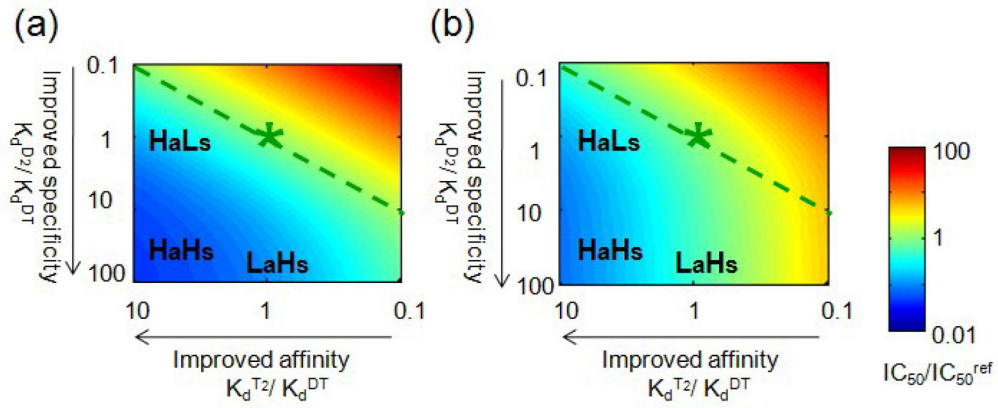

Figure 7.

Inhibition of DNA binding as a function of the affinity and anti-homodimer specificity of the inhibitor. A description of the model is given in Methods. The ratio of the IC_{50} for a design to the IC_{50} for a reference inhibitor with affinity equal to the wild-type protein is used as an indicator of design potency (scale at right). This ratio is plotted as a function of the affinity and specificity of the inhibitor. In (a), the K_d values for target dimerization and DNA binding are 10-fold lower than the bZIP concentration. In (b) the K_d values for both associations are 10-fold higher than the bZIP concentration. Labeling on the graph (H_aL_s $H_{affinity}L_{spec}$, L_aH_s: $L_{affinity}H_{spec}$ and H_aH_s: $H_{affinity}L_{spec}$) is described in Discussion. The dashed line represents designs with zero hetero-specificity. The reference inhibitor is indicated with a star.

Table 1

Sequences[a] and melting temperatures (°C)[b] for BZLF1 and design constructs.

Constructs	basic/acid	coiled coil 191 221 bcdefgabcdefgabcdefgabcdefgabcd	proximal CT 231	distal CT 245	T_m
BZLF1^{231}		AKFKQLLQHYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII		43
A-BZLF1^{231}	QRAEELARENEELEKEA	EELEQELLKYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII		33
B-BZLF1^{231}	LEIKRYKNRVASRKCR	AKFKQLLQHYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII		31
BZLF 1^{245}		AKFKQLLQHYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	71
A-BZLF1^{245}	QRAEELARENEELEKEA	EELEQELLKYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	43
B-BZLF1^{245}	LEIKRYKNRVASRKCR	AKFKQLLQHYREVAAAKSSENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	67
BD_{CC}^{231}		AKEEQEIQHLEEEIAALESENDRLRLLLKQM	CPSLDVDSII		38
BD_{CC}^{245}		AKEEQEIQHLEEEIAALESENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	40
$A-BD_{CC}^{245}$	QRAEELARENEELEKEA	EELEQELLKLEEEIAALESENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	40
BD_{IED}^{231}		AKFKQLLQHYREVIAAEDSENDRLRLLLKQM	CPSLDVDSII		N/A[c]
BD_{IED}^{245}		AKFKQLLQHYREVIAAEDSENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	26
$A-BD_{IED}^{245}$	QRAEELARENEELEKEA	EELEQELLKYREVIAAEDSENDRLRLLLKQM	CPSLDVDSII	PRTPDVLHEDLLNF	N/A[c]
BD_{EEI}^{231}		AKFKQLLQHEEEVIAAKSSENDRLRLLLKQM	CPSLDVDSII		N/A[c]

[a] The sequences SHHHHHHGESKEYKKGSGS, or GYHHHHHHGSY (the latter for constructs with the acidic extension, A-) should be placed at each N terminus to obtain the full sequences of the recombinant proteins listed in the table. Sites with amino acids different from those of the native sequence (either introduced in the design or as part of the acidic extension) are underlined. Different regions of the sequence (basic region/acidic extension, coiled coil, proximal CT and distal CT) are separated by space. As explained in the text, the acidic extension overlaps the 9 N-terminal residues of the coiled coil. Coiled-coil heptads are indicated using shading.

[b] Total protein concentration was 4 μM.

[c] N/A indicates either lack of cooperative folding or that the observed melting curve indicated the presence of more than one species.

Table 2

Melting temperatures (°C) for different BZLF1/design hetero-interactions.

Target	Design	T_m [a]	ΔT_m [b]
$BZLF1^{231}$	BD^{231}_{CC}	53	12 (43/38)
	BD^{231}_{IED}	N/A[c]	N/A[c]
$B\text{-}BZLF1^{245}$	BD^{245}_{CC}	66	12 (67/40)
	BD^{245}_{IED}	N/A[c]	N/A[c]
	$A-BD^{245}_{CC}$	>80	> 26 (67/40)
	$A\text{-}BZLF1^{245}$	74	19 (67/43)
$B\text{-}BZLF1^{231}$	$A\text{-}BZLF1^{231}$	58	26 (31/33)
JUN	BD^{245}_{CC}	41	10 (23/40)

[a] Total protein concentration was 4 μM.

[b] ΔT_m was obtained by taking the T_m for the hetero-complex and subtracting from it the average of the T_m values for each individual species (listed in parentheses for easy comparison, T_m for the target is shown first, followed by that of the design) when applicable.

[c] N/A indicates either lack of cooperative folding or that the observed melting curve indicated the presence of more than one species.