Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design

Jody M. Mason, Mark A. Schmitz, Kristian M. Müller, and Katja M. Arndt*

Institut für Biologie III, Universität Freiburg, Schänzlestrasse 1, D-79104 Freiburg, Germany

Edited by Jennifer A. Doudna, University of California, Berkeley, CA, and approved April 18, 2006 (received for review November 14, 2005)

Activator protein-1 (AP-1) is a crucial transcription factor implicated in numerous cancers. For this reason, nine homologues of the AP-1 leucine zipper region have been characterized: Fos (c-Fos, FosB, Fra1, and Fra2), Jun (c-Jun, JunB, and JunD), and semitational library-designed winning peptides FosW and JunW. The latter two were designed to specifically target c-Fos or c-Jun. They have been identified by using protein-fragment complementation assays combined with growth competition. This assay removes nonspecific, unstable, and protease susceptible library members from the pool, leaving winners with excellent drug potential. Thermal melts of all 45 possible dimeric interactions have been surveyed, with the FosW–c-Jun complex displaying a melting temperature (Tm) of 63°C, compared to only 16°C for wild-type c-Fos–c-Jun interaction. This impressive 70,000-fold ΔTm decrease is largely due to optimized core packing, α-helical propensity, and electrostatics. Contrarily, due to a poor c-Fos core, c-Fos–JunW dimerizes with lower affinity. However the Tm far exceeds wild-type c-Fos–c-Jun and averaged JunW and c-Fos, indicating a preference over either homodimer. Finally, and with wider implications, we have compiled a method for predicting interaction of parallel, dimeric coiled coils, using our Tm data as a training set, and applying it to 59 bZIP proteins previously reported. Our algorithm, unlike others to date, accounts for helix propensity, which is found to be integral in coiled coil stability. Indeed, in applying the algorithm to these 59 bZIP interactions, we were able to correctly identify 92% of all strong interactions and 92% of all noninteracting pairs.

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: AP-1, activator protein-1; CC, coiled coil.

*To whom correspondence should be addressed. E-mail: katja@biologie.uni-freiburg.de.

© 2006 by The National Academy of Sciences of the USA
Results

Targeting c-Fos: Design of Jun Library and Selection of JunW. Despite being a reasonable CC, with good core and electrostatic properties, the Jun interaction pattern with Fos is still taxing to understand. Consequently, in library design, wild-type core residues have been retained, with similar options introduced (Fig. 1A). Specifically, three β-branched residues (I, V, T), and A at α1, α2, α4, and α5, were added (12, 13). Mutation of α1 to L has previously yielded remarkably increased stability for a dominant negative Jun–Fos (14). However, in the interest of library size, and additional undesirable library options, L was excluded. Interestingly, of these four varied positions, A arose twice in the selection, at α1 and α4. V at α2 was unchanged, whereas α5 switched from V to I. Selection of A at positions α1 and α4 were late in growth competitions, indicating low preference. Indeed, the JunW core is little improved over parent molecule or homologues.

The N–K α3 pair observed in the Jun-W heterodimer was kept, along with I and K options in the library. Despite instability incurred from N–N and K–K pairings, hydrogen bonding plays a role in specificity determination (15). This has been disputed on the grounds that both a–g (parallel) and a–e (antiparallel) K–E interactions can occur, with K side chain internal methylene groups maintaining core integrity (16). Additionally, α1–α2 pairs are stable compared to N–K, although specificity conferred by such pairings are unknown. Regardless, the option is a good test-bed within our experimental confines, as either c-Fos–c-Fos or JunW–JunW homodimers forming preferentially over c-Fos–JunW will slow bacterial growth and be removed from the pool. Interestingly, in the selection process N and K (opposite c-Fos K) were rejected from α3 in favor of I.

In parallel dimeric CCs, a g residue can form a coulombic interaction with an e i+1 residue of the next heptad on the opposite helix (6). In the JunW, the e3 position was changed from A to Q with both R and K rejected (Fig. 1A). All replacements were predicted to pair well with g2 E in c-Fos (17). All c-Jun g positions have been retained in the library, with alternatives aiming for equivalent or better electrostatic interactions with c-Fos e residues. However, no wild-type residues were selected and, excepting the α1 N-cap motif, no other changes were made to the molecule owing to library size and a lack of suitable homologue variations.

Positions g1, g2, and α2, settled fastest, and are perhaps pertinent in heterodimer formation or stability. g1 R was predicted to pair well with c-Fos c2 E while shielding the core better than alternatives Q or K. Both α1 and α4 fluctuated between V and A, with the latter selected last, suggesting no major preference between these residues at this position.

Targeting c-Jun: Fos Library Design and Selection of FosW. In core design, Fos has much scope for improvement comparative to Jun. Consequently, FosW–c-Jun will always be more stable than a corresponding JunW–c-Fos. The instability of c-Fos (it cannot homodimerize) is the principal heterodimeric driving force for c-Jun–c-Fos formation, rather than heterodimeric preference over c-Jun–c-Jun homodimers (18). Repulsive g/e interactions and poor core α residues largely explain this phenomenon (Fig. 1B). For example, K in the Fos core (and E at g1/e2) helps prevent Fos homodimerization (19). Two α positions K and E residues, conserved in Fos homologues, were proposed to be central in accounting for most additional heterodimerization. Consequently, Fos library wild-type α positions were removed and replaced with β- and γ-branched options (L, V, I), for better packing and desolvation (Fig. 1B). In selection, L or I, but not V arose in all four instances. Position α3, opposing c-Jun N, was fixed from K to N; this was predicted to be the best pairing for specificity and to be favored over aliphatic–polar combinations (12).

g/e i+1 interacting pairs, which complement c-Jun, were introduced according to Vinson’s free energy values (17). g1 and c2 E residues, conserved in Fos homologues, were proposed to be central in accounting for most additional heterodimerization free energy with c-Jun (22), and were not varied in the library. Variations from Fos homologues were included in library design. In winning g and e positions, no wild-type residues were selected. Other options such as e3 R were reasoned to form g/e i+1 interactions with g2 Q of c-Jun. Some changes (all from homologue variations, see Fig. 1B) are difficult to rationalize because they deviate from classical charged/polar pairings. For example, in FosW g2 R pairs with c-Jun e3 A. Other changes, often subtle (e.g., F2 D → E, found in FosB and Fra2; C4 N → E, found in FosB, Fra2 and dFra, the latter being D in our library for coding reasons), appear to play no direct interfacial role. Strikingly, all homologue residues were selected in FosW, and collectively these changes introduce a higher extent of polar residues, with little overall change in pI.
burial, propensity, solubility, electrostatic attraction of flanking residues, and a number of intramolecular interactions. Ranking contributions, and building a relationship between residue changes and $T_m$, is consequently very difficult. However, three overriding factors (core, electrostatic and propensity) have been considered. Accordingly, we have devised an improved algorithm for CC prediction using core, electrostatic, as well as propensity. This algorithm can be found online at www.molbiotech.uni-freiburg.de/bCIPA and is known as the bZIP coiled coil interaction prediction algorithm (bCPIPA).

A rudimentary core packing score has been assigned to all dimers to distinguish cores which make large contributions to stability from those which do not (Fig. 3) by scoring hydrophobic pairings highest, with $aa'$ and $dd'$ pairs treated the same for simplicity. $LL$ ($LL = −1.5$) is ranked higher than all others owing to its exclusivity in d positions. Other hydrophobic pairings ($VV = II = LV = L1 = VI = −1$) as well as $KI$ and $NN$ combinations ($KI = NN = −1$) were favored over other lysine and alanine combinations ($KL = KV = LA = VA = IA = IT = LT = AA = −0.5$) with a further subset being disfavored entirely ($IN = LN = VN = TV = +0.5$). These rankings are based on $\Delta G$ energies from $aa'$ pairings relative to $AA$ pairs (23), observance in other CC proteins, and our own unpublished results. No preference was given for homotypic or heterotypic hydrophobic core pairing. $KI$ has been described to be more stable than $KV$ and $KL$, for reasons that are not yet understood (23), and also $KI$ was selected twice in c-Fos–JunW. Finally, $NN$ was favored significantly over $KK$, possibly because of increased steric restraint and charge repulsion in the latter. Stability offered by $NN$ pairing is relatively low. $N$ has low propensity and high polarity for a core region, but confers specificity by limiting oligomeric states to dimers (12), this benefit outweighing lack of stability.

Our electrostatic parameters are based on opposing charge pairings and place energetic penalties on similar charge pairings ($DD = DE = EE = RR = KR = −1; KD = RD = EQ = −0.5; KQ = RQ = −1; QO = KE = RE = −1.5$) with $gi/e_{i+1}$ and $e_{i+1}/g_i$ interactions treated the same for simplicity. Electrostatic interactions were related to free energy contributions based on data from a double mutant analysis (17). Consequently, although not included implicitly in the study, the scale indicates much improved electrostatic attractions.

Fong et al. (24) used “base optimized weighting” to predict CC interactions, and identified strong interactions based on $dd'$, $aa'$, $ad'$, $da'$, $de'$, $ga'$, and $ge'$ pairings (24), but did not consider $aa'$-helical stability as a direct contributing factor. However, we estimate helical propensity to be hugely important and a largely overlooked third parameter in CC stability (covered in depth in Discussion), precluding electrostatic and core considerations in forming a structure which is in a dimerization competent state. Surprisingly, we find that only two of the seven considerations made by Fong et al. (24) are strictly necessary ($da'$, $da'$, $de'$, $ga'$ pairings are not required, and $aa'$ and $dd'$ count as one), and that propensity is a more important omission. Indeed, the role of surface residues has been probed for GCN4-p1, with helix propensity found to be a key factor in surface design (25). Additionally, intramolecular hydrogen bonding of high propensity residues such as $Q$, $R$, $E$, or $K$, which frequent these positions is also important, and if unsatisfied, can cause unfavorable effects. In combining these parameters with a least squares fit, stability can be rationalized and agreed well with actual $T_m$ values (Fig. 4). It should be noted that our algorithm fits $T_m$ values that indicate pairing stabilities. For specificities ($\Delta T_m$), the propensity term cancels out (see also supporting information, which is published on the PNAS web site).

In our fitting procedure, core pairings ($dd' = a/a'$), $g/e$ electrostatic preferences and propensity scales from Williams et al. (26) were used to fit our $T_m$ data (45 dimers) as well as selected and rationally designed peptides characterized previ-

In the only comprehensive, semiquantitative, study to date (27), Jun–Fos homologues range from 84–107 residues, contain terminal sublining sequences, a his-tag, an N-terminal basic region extension, and at least nine residues of C-terminal extension until the paircoil program (37) predicts the probability of CC to be <10%. In contrast, our peptides are all 37-mers of same register from the CC region, N- and C-capped, and are not disulfide bridged. Despite these differences, the core CC sequences (excepting a b position Y for absorbance) are identical in both studies, and display moderate agreement with our own \( T_m \) values. Using Z scoring (27) as an affinity measure, that earlier study showed in a microarray analysis heterodimers to yield a good interaction (Z > 10). Homodimeric Fos typically displayed low stabilities, and homodimeric Jun combinations (excepting c-Jun; \( Z > 5 \), \( T_m \) 40°C) were less stable than heterodimers. We found c-Jun–c-Jun to be more stable than c-Jun–c-Fos; however, our data do not dispute that c-Fos instability drives heterodimeric preference (18). Regardless, excluding electrostatic considerations, c-Jun–c-Jun has better \( \alpha \)-helical propensity, and a significantly more stable core than c-Fos–c-Jun. Importantly, to our knowledge, our study is the first quantitative biochemical analysis of all possible Jun-Fos CC combinations from human (see Fig. 3).

Core Conclusions from Winners. Core comparison indicates that dimeric Fos homologue cores are disfavored compared to Fos–Jun, with Jun cores displaying greater hydrophobic burial and greater stability. Winners have similar (JunW) or enhanced (FosW) cores compared to homologues, resulting in optimized cores and improved stabilities.

In general, I is favored over V in winner a positions, has higher propensity, and is bulkier for core packing. It has been reported to be more stable and better than V and L at this position in conferring dimers (12, 13). FosW L and I were selected over V in all four instances, despite a documented B- over \( \gamma \)-branching preference (12, 38). V alone yields trimers (39) or a mixture of dimers and trimers (12), whereas I can specify dimers exclusively.

Surprisingly, no homotypic pairing preference was observed in FosW–c-Jun, with aa’, LI, IV, IA, and LV selected. Nor were any \( a1 \) or \( a2 \) TT pairings observed despite speculation that generally homotypic pairings are energetically favored over similar hydrophobic pairings (23). In contrast, the bulk of core winner aa’ pairings are heterotypic.

g/e1+,c Conclusions from Winners. Abundance of strong interacting pairs involving (with the exception of Q) terminal charge attractions such as KE, RE, QE, QQ, RO, and KQ, suggest that hydrophobic bulk plays an additional role (40, 41). The side chain of E (−(\( CH_2 \))_2COO\(^- \)) can pair with K (−(\( CH_2 \))_3=NH\(^+ \)) and R (−(\( CH_2 \))_3=NH–CN\( (NH)(NH)_2 \)) side chains, both having positively charged termini able to contact the negative carboxyl group of E. Q is also favorable, possibly because its side chain (−(\( CH_2 \))_3CONH\(^+ \)) is of sufficient length to shield the core from the solvent. However, lack of terminal charge is predicted to lower the specificity of the interaction. Partial hydrophobic environment felt by the side chains may also improve the energetic contribution of these charge interactions, yielding a greater contribution in less aqueous surroundings.

However, from frequency in CCs (40, 41) and energetic rankings (17), we conclude that D and N (both of which are shorter, containing only one side chain methylene group) are disfavored and should be omitted in designed coils, despite possible charge complementarity (for D) with K or R again due to poor core shielding.

The extra length and polarity of the JunW e3 A → Q change is likely to enable contact with the corresponding E at c-Fos g2.
Additionally, the extra hydrophobic bulk of Q’s β- and γ-carbon methylene groups can stabilize the molecule by shielding the core from the solvent to a greater extent than can A (4), and these methylene groups may provide favorable interactions with the δ-carbon methylene groups of c-Fos L. Electrostatic attractions estimated according to Krylov et al. (17) in c-Fos–JunW (ΔΔG = −5.2 kcal/mol), are improved compared to the c-Fos homodimer (−0.6 kcal/mol) and surpass the average of c-Fos and JunW homodimers (−2.4 kcal/mol), indicating an electrostatic preference for heterodimer formation.

**Outer Positions: Intrahelical Stability and Solubility.** Incorporating Fos homologue residue changes at solvent exposed regions interestingly resulted in acceptance of all proposed amino acids at all nine positions in the library winner (Fig. 1). Depending on the scale used, either five (26) or six (42) of these changes were found to have a higher helical propensity than wild-type c-Fos. FosW A → E change at position b4 could be interacting favorably with Q at c4 and K at f4, with the resulting increase in helix stability propagating to stabilize the CC. Other solvent exposed changes such as T → K, N → D, and D → E, may act by contributing to increased helical propensity, with charges increasing protein solubility, and aiding the driving force for folding by increasing the concentration of monomers in a dimerization competent state.

**Helical Propensity Considerations.** Propensity scales inform upon the frequency or preference with which a given residue occurs in a particular conformation. In our analysis we have used the scales devised by Williams et al. (26) as well as Gromiha and Parry (42). From these scales, averaged helical propensity predictions were assigned to each of the helices, discounting N- and C-caps. Williams et al. is similar to the Chou and Fasman scale (43) in that it is derived from statistical data, whereas Gromiha and Parry is derived specifically from CCs, with the former scale giving a mildly better fit to our data set. We favored these scales over other experimental approaches (44–46) because they include not only substituted solvent exposed residues at the center of a helix, but partially and completely buried residues together with residues at the helix termini, both of which would differ in propensity in these contexts (47, 48). This finding is of particular significance in short helical CC motifs such as ours, where residues are completely buried, partially buried, or completely solvent exposed, depending on side chain and heptad position, or centrally or terminally located. Analysis of the homologues predicts winning peptides will display increased helicity. This is somewhat surprising given that it was not a criterion in library design. However, homologues are informative in this respect because they contain higher proportions of the destabilizing design. However, homologues are informative in this respect somewhat surprising given that it was not a criterion in library design. FosW A → E change at position b4 could be interacting favorably with Q at c4 and K at f4, with the resulting increase in helix stability propagating to stabilize the CC. Other solvent exposed changes such as T → K, N → D, and D → E, may act by contributing to increased helical propensity, with charges increasing protein solubility, and aiding the driving force for folding by increasing the concentration of monomers in a dimerization competent state.

No specific sequences are identified that conform to a speculated “trigger sequence” (50), although propensity may play a crucial role during folding, possibly acting to enforce α-helical topology, thus ensuring structures are driven thermodynamically and on a biologically realistic time scale. FosW and JunW contain no central G residues that are replaced with higher propensity residues. Closely connected, but much harder to predict, is the role of context in determining stability, or how interactions of side chains with surrounding side chains affect overall propensity.

In the future, it should be possible to factor in a greater negative design aspect where design for specificity, as well as stability, can play an increased role. More difficult will be designing to generate the lowest possible Kd, while retaining specificity so that requirements of the proteins are met. All these considerations must be accounted for as well as being incorporated into library designs, to design coils that are both stable and specific.

**CC Prediction.** We have rationalized our winning peptides based upon well understood principles. The algorithm of Fong et al. (24) fails noticeably short in predicting the Tm values for Fos–Fos homologues (Fig. 4B). Our own algorithm is much improved with respect to this subset. Additionally, the Fong algorithm is less able to predict Tm values of heterodimeric winner complexes, whereas ours underestimates only three winner complexes, notably c-Jun–JunW, JunD–JunW, and JunB–JunW, but makes reasonable estimates of the remainder.

We have used a combination of simplistic core and electrostatic parameters, combined with well documented but little implemented (in CC stability prediction) helical propensity scales. In combining these parameters, we have devised a prediction algorithm. Although this has further potential for optimization, it is at least in the context of bZIPs on par with and certainly more simplistic than that of Fong et al. (24). A corollary of this work is to design more robust CC pairs and dominant negatives with improved therapeutic value, and as a potential use as building blocks in nanobiotechnological design.

**Materials and Methods**

**Library Design and Cloning.** Mega-primers were synthesized including relevant degenerate codons for residue options (for libraries), and a fill-in reaction was performed, resulting in 111-bp double-stranded oligonucleotides. These were cloned via...
Nhe1 and Ascl sites into a pQE16 derivative (Qiagen) containing a G/S linker tagged to fragment 1 (pAR200d; c-Jun and Jun library; ampicillin resistance; K.M.A., unpublished data) or fragment 2 (pAR300d; c-Fos and Fos library; chloramphenicol resistance; K.M.A., unpublished data) of murine dihydrofolate reductase (mDHFR), respectively. Library plasmids were transformed into BL21 gold cells (Stratagene) containing target plasmid and pREP4 (Qiagen; for lac repression). To assess library quality, we sequenced pools and single clones and found approximately equal distributions of varied amino acids. Pooled colonies exceeded the library size 5- to 10-fold.

**Selection of Winner Peptides.** The protein-fragment complementation assay has been described (8, 9, 11). Briefly, CCs are tagged to either half of murine dihydrofolate reductase. Only two interacting helices will bring the two halves of the enzyme into close proximity, render the enzyme active, and result in colony formation on M9 minimal medium plates with trimethoprim (1 μg/ml) and tetracycline (10 μg/ml). Selection on M9 plates supplemented with 1 mM K-phosphate (pH 7) using a Jasco J-810 CD instrument. The temperature was ramped at a rate of 0.5°C per min. Melting profiles were ≥94% reversible with equilibrium denaturation curves fitted to a two-state model to yield the melting temperature (Tm).

\[ \Delta G = (\Delta H - T \Delta S) = (\Delta H + R \times T \times \ln(P_i)) + \Delta C_p \times \left( T_m - T - \ln \left( T_m/T_m^0 \right) \right), \]

where \( \Delta H \) is the change in enthalpy, \( T \) is the reference temperature; \( R \) is the ideal gas constant; \( P_i \) is the total peptide concentration; and \( \Delta C_p \) the change in heat capacity. Additionally, \( T_m \) of a heterodimer AB is calculated by using \( \Delta T_m(AB) = T_m(AB) - 0.5(T_m(A) + T_m(B)) \).

**Stability Prediction.** Helix propensity (HP) is calculated as an average over the whole helix, i.e., the individual residues are summed divided by the total number of residues. Electrostatic (ES) and core (C) are calculated by using a simple weighting scheme (see Results) and summed over the whole peptide to account for increased stability in longer helices. Scores for measured \( T_m \) values were fitted as follows

\[ T_m = a_1 \times HP + a_2 \times C + a_3 \times ES + d, \]

where \( a_1, a_2, \) and \( a_3 \) are weighting factors for the three parameters, and \( d \) is an offset factor. Temperatures were fitted in Kelvin (see www.molbiotech.uni-freiburg.de/bCIPA).

More detailed descriptions can be found in supporting information.

This work was funded by the Emmy Noether program of the Deutsche Forschungsgemeinschaft (Grant Ar373).