

A Heterodimeric Coiled-coil Peptide Pair Selected *in Vivo* from a Designed Library-versus-Library Ensemble

Katja M. Arndt¹, Joelle N. Pelletier^{1,2}, Kristian M. Müller¹, Tom Alber³
Stephen W. Michnick² and Andreas Plückthun^{1*}

¹Biochemisches Institut
Universität Zürich
Winterthurerstr. 190
CH-8057, Zürich, Switzerland

²Département de biochimie
Université de Montréal
Montréal, Québec, Canada
H3T 1J4

³Department of Molecular and
Cell Biology, University of
California, Berkeley, CA
94720-3206, USA

Novel heterodimeric coiled-coil pairs were selected simultaneously from two DNA libraries using an *in vivo* protein-fragment complementation assay with dihydrofolate reductase, and the best pair was biophysically characterized. We randomized the interface-flanking e and g positions to Gln, Glu, Arg or Lys, and the core a position to Asn or Val in both helices simultaneously, using trinucleotide codons in DNA synthesis. Selection cycles with three different stringencies yielded sets of coiled-coil pairs, of which 80 clones were statistically analyzed. Thereby, properties most crucial for successful heterodimerization could be distinguished from those mediating more subtle optimization. A strong bias towards an Asn pair in the core a position indicated selection for structural uniqueness, and a reduction of charge repulsions at the e/g positions indicated selection for stability. Increased stringency led to additional selection for heterospecificity by destabilizing the respective homodimers. Interestingly, the best heterodimers did not contain exclusively complementary charges. The dominant pair, WinZip-A1B1, proved to be at least as stable *in vitro* as naturally occurring coiled coils, and was shown to be dimeric and highly heterospecific with a K_D of approximately 24 nM. As a result of having been selected *in vivo* it possesses all characteristics required for a general *in vivo* heterodimerization module. The combination of rational library design and *in vivo* selection presented here is a very powerful strategy for protein design, and it can reveal new structural relationships.

© 2000 Academic Press

Keywords: coiled coil; leucine zipper; PCA; protein design; protein engineering

*Corresponding author

Introduction

α -Helical coiled coils are involved in the oligomerization of a wide variety of proteins. Because of their small size and structural regularity, they have also been used as artificial domains to med-

iate oligomerization of various proteins (Pack & Plückthun, 1992; Pack *et al.*, 1995; Weissenhorn *et al.*, 1997). Coiled coils consist of two or more amphipathic helices wrapping around each other with a slight supercoil. They contain a characteristic heptad repeat $(a-b-c-d-e-f-g)_n$ with a distinct pattern of hydrophobic and hydrophilic residues (Figure 1(a), for reviews see Hurst, 1995; Hodges, 1996). The positions a and d, which form the hydrophobic interface between the helices, are usually aliphatic and have a profound effect on the oligomerization state (Harbury *et al.*, 1993, 1994). The positions b, c, e, g, and f are solvent-exposed and usually polar. The positions e and g, which flank the hydrophobic core, can make interhelical interactions between g_i and e'_{i+5} residues, and thereby mediate heterospecific pairing (Graddis

Present addresses: K. M. Arndt and K. M. Müller,
Department of Molecular and Cell Biology, University
of California, Berkeley, CA 94720-3206, USA.

Abbreviations used: CD, circular dichroism; DHFR,
dihydrofolate reductase; WinZip-A1, predominantly
selected peptide from library A; WinZip-B1,
predominantly selected peptide from library B; WinZip-
A1B1, heterodimer of the peptides WinZip-A1 and
WinZip-B1.

E-mail address of the corresponding author:
plueckthun@biocfebs.unizh.ch

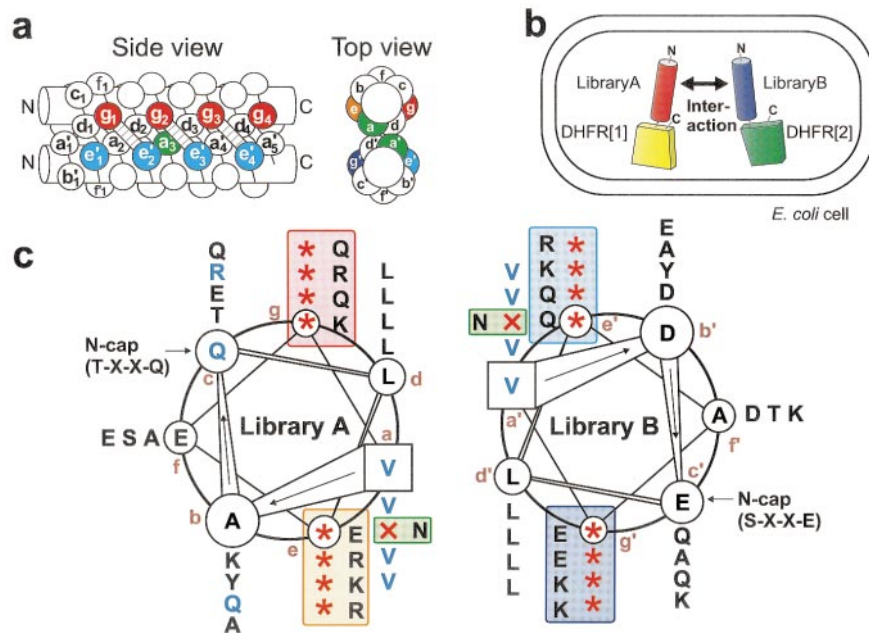


Figure 1. (a) Schematic representation of a parallel dimeric coiled coil. Side view: the helical backbones are represented by cylinders, the side-chains by knobs. The path of the polypeptide chain is indicated by a line wrapped around the cylinders. For simplicity, the supercoiling of the helices is not shown. While residues at positions **a** and **d** make up the hydrophobic interface, residues at positions **e** and **g** pack against the hydrophobic core. They can participate in interhelical electrostatic interactions between residue i (**g** position) of one helix and residue $i' + 5$ of the other helix (e' position, belonging to the next heptad), as indicated by the hatched bars. Top view: arrangement of the heptad positions. (b) Schematic representation of the protein-fragment complementation assay as described in the text. Each library is genetically fused to one of the two DHFR-fragments ([1] or [2]), and only an interaction between the two library peptides restores the enzyme activity which is crucial for cell survival under selective conditions. Introduction of a mutation at the DHFR interface (I114A) was used to increase selection stringency where indicated (Pelletier *et al.*, 1998). (c) Overview of the library design depicted as a helical wheel plot from the N to the C terminus (inside to outside). The black residues correspond to the original residues from c-Jun (library A) and c-Fos (library B). Changes introduced by the design are in blue. The randomized positions are in red (*, equimolar mixture of Q, E, K, R; ×, equimolar mixture of N, V) and are boxed with the same colors as used in (a). The selected residues from the predominant pair, WinZip-A1B1 (clone 1 in Table 3), are next to the randomized positions in the respective boxes.

et al., 1993; O'Shea *et al.*, 1993; Zhou *et al.*, 1994; Kohn *et al.*, 1995; Jelesarov & Bosshard, 1996).

As most naturally occurring coiled coils are homodimeric, synthetic sequences have been designed to promote specific hetero-oligomerization (Graddis *et al.*, 1993; O'Shea *et al.*, 1993; Nautiyal *et al.*, 1995; Jelesarov & Bosshard, 1996). Although progress in understanding many properties, such as stability and heterospecificity, has been made by cycles of rational design and analysis, several questions are still remaining which cannot be solved by such an approach. Most importantly, these concern the expression properties in hosts such as *Escherichia coli* and the resistance against proteases. We previously observed that designed coiled coils which behave well as synthetic peptides failed in fusion proteins expressed in *E. coli*, as they were proteolytically degraded (K.M.A., P. Pack & A.P., unpublished results). An interesting application for *in-vivo* selected heterodimeric coiled coils would be the generation of bispecific miniantibodies (Plückthun & Pack, 1997; Müller *et al.*, 1998a,b).

To address this challenge we combined the knowledge of various existing rational designs with the advantages of library diversity. This presents a formidable challenge to most of the well-known selection systems such as phage, bacterial or ribosome-display, as in all these cases a direct selection of two simultaneously varied, unlinked proteins is not straightforward or not possible at all. While the yeast two-hybrid system might in principle be used for this purpose, bacterial expression gives, in general, access to larger libraries because of their more efficient transformation. The λ -repressor dimerization strategy, which has been applied to a similar problem (Hu *et al.*, 1993; Zeng *et al.*, 1997), requires additional complex screening and interpretation to distinguish between homo- and heteromerization as well as between dimers and higher-order oligomers. We therefore took advantage of a novel selection system based on protein-fragment complementation that allows simultaneous selection for heterodimers from two libraries in *E. coli* (Pelletier *et al.*, 1998, 1999) (Figure 1(b)). Each library is genetically fused to one of two designed fragments of the enzyme

murine dihydrofolate reductase (mDHFR). Only cognate interaction between members of both libraries can reassemble the functional recombinant murine DHFR. This activity is essential for propagation of *E. coli* in minimal medium in the presence of trimethoprim, which selectively inhibits the prokaryotic DHFR.

By applying three different stringencies of selection, three sets of selected sequence pairs were obtained, and the most stringent selection was dominated by one pair, named WinZip-A1B1 (Pelletier *et al.*, 1999). Sequence comparison revealed a very strong selection for Asn pairs in the core **a** position over Val-Val or Val-Asn combinations, suggesting a parallel, dimeric association of the helices. An increase of opposite-charged interactions and a decrease of same-charged interactions in the **e/g** pairing of the populations obtained, which on the whole is consistent with the current view of coiled coils, led us to conclude that we had indeed selected for heterodimeric coiled coils. However, this still remained to be proven by biophysical analysis of the dominant pair, WinZip-A1B1. Furthermore, we were still lacking information on whether the selection system was able to not only select for stability but also for heterospecificity.

Here, we describe a detailed analysis of pairs obtained from the three different stringency selections and a biophysical characterization of the selected heterodimeric peptide in comparison to the homodimers of its constituting peptides. This has enabled us to rank positions and types of interhelical interactions with respect to their importance, revealing deeper insight into the selection process and into coiled-coil interactions. Interestingly, the selected pairs revealed a more complex interaction pattern than used in rationally designed coiled coils. The predominant pair WinZip-A1B1 from the most stringent selection was characterized as synthetic peptides with a series of biophysical techniques. These *in vitro* data are in excellent agreement with the trends extracted from the sequence analysis of the *in vivo* selections and led us to conclude that we have obtained a novel heterospecific and metabolically stable coiled coil.

Results

Our goals were to understand the driving forces of heterospecific coiled-coil interactions and ultimately, to select for a metabolically stable dimeric coiled coil with high heterospecificity. Two libraries were designed to meet the requirements of genetic diversity to prevent recombination, high helix stability and a high probability of complementarity, all within a reasonable library size (Figure 1(c)). While the libraries had been used in a previous proof-of-principle of the selection system (Pelletier *et al.*, 1999), we focus here on issues of the library design which are essential for under-

standing the sequence and biophysical analysis of the selected pairs.

Library design

As templates for the outer, solvent-exposed residues (positions **b**, **c**, **f**) we chose the leucine-zipper regions of the proto-oncogenes c-Jun and c-Fos for library A and library B, respectively. This study minimized potential recombination despite the repetitive pattern of the heptads (for reviews of c-Jun and c-Fos, see Curran & Franza, 1988; Ransone & Verma, 1990). Indeed, no recombination was found in any of the 80 clones sequenced. We chose a helix length of 4.5 heptads as a good compromise between stability and size. For the hydrophobic core residues (positions **a**, **d**) we chose the residues of the parallel, homodimeric leucine zipper GCN4 (Val at **a**, Leu at **d**). These are commonly found in all coiled coils. A single **a** position in the middle of each helix is often occupied by a polar residue, most often Asn, which forms a hydrogen bond inside the hydrophobic core (O'Shea *et al.*, 1991; Junius *et al.*, 1995). Replacement of this Asn pair by a non-polar one increases the stability significantly, but leads to helices packing in different registers and orientations, as well as forming higher-order oligomers (Potekhin *et al.*, 1994; Lumb & Kim, 1995; Ogihara *et al.*, 1997). Since we could not ascertain *a priori* whether higher specificity or stability would be more advantageous, we included both by allowing Asn and Val at the core **a** position with equal probability. A difficult problem in library design is to encode only the desired amino acids with a predetermined ratio. We solved this problem by using defined trinucleotide mixtures in the oligonucleotides, where each trinucleotide codes for one specific amino acid (Virnekäs *et al.*, 1994).

The solvent-accessible residues at the **e** and **g** positions can form interhelical salt bridges or hydrogen bonds which can contribute to stability and heteromeric specificity (O'Shea *et al.*, 1993; John *et al.*, 1994; Krylov *et al.*, 1994; Zhou *et al.*, 1994). Based on these results and on commonly occurring amino acids at these positions (Cohen & Parry, 1990; Woolfson & Alber, 1995), we simultaneously randomized all eight **e** and **g** positions with equimolar mixtures of Gln, Glu, Lys, Arg, also using trinucleotide codons in DNA synthesis. Including the Asn-Val combination at the core **a** position, each library had a theoretical diversity of 1.3×10^5 .

To increase the stability of the helices, we introduced helix capping residues in both libraries to saturate the missing hydrogen bonds at the helix ends with their side-chains. Based on studies of helix capping propensities in proteins (Richardson & Richardson, 1988; Dasgupta & Bell, 1993) and peptides (Chakrabarty *et al.*, 1993; Doig & Baldwin, 1995), we chose T-X-X-Q (N-cap-N1-N2-N3) for library A, and S-X-X-E for library B. In both cases, the N-cap side-chain (T, S, respectively) should form a hydrogen bond with the N3 amide

Table 1. Sequence pairs selected in the lowest stringency selection (single-step selection using wild-type DHFR fragments)

| Clone | Library A ^a | | | | | | | | | Library B ^b | | | | | | | | |
|-------|------------------------|----|----|----|----|----|----|----|----|------------------------|-----|-----|-----|-----|-----|-----|-----|------|
| | e1 | g1 | e2 | g2 | a3 | e3 | g3 | e4 | g4 | e'1 | g'1 | e'2 | g'2 | a'3 | e'3 | g'3 | e'4 | g'4 |
| 1 | R | K | Q | E | V | E | K | E | E | Q | Q | K | E | V | R | R | K | K |
| 2 | K | Q | Q | E | N | Q | K | R | K | K | E | Q | R | N | K | K | E | R |
| 3 | R | Q | R | Q | N | Q | R | Q | K | Q | Q | Q | R | V | Q | Q | E | K |
| 4 | E | Q | R | E | V | E | Q | K | R | Q | Q | Q | R | V | K | E | E | K |
| 5 | R | Q | Q | K | N | Q | Q | Q | E | E | Q | K | K | N | R | Q | Q | R |
| 6 | E | E | K | Q | N | R | R | Q | R | Q | E | R | R | N | R | Q | Q | K |
| 7 | E | Q | E | K | N | E | E | K | E | E | K | R | K | V | E | R | K | K |
| 8 | K | K | K | Q | N | E | K | K | R | K | Q | K | E | N | K | R | E | Q |
| 9 | E | Q | Q | Q | V | Q | R | R | R | E | K | K | K | V | Q | R | Q | R |
| 10 | K | R | E | Q | N | Q | R | R | E | K | K | R | H | N | R | E | E | R |
| 11 | Q | Q | R | Q | N | Q | E | K | K | Q | K | R | Q | N | K | E | K | Q |
| 12 | Q | E | K | Q | N | E | K | E | R | K | Q | K | R | N | R | R | Q | Q |
| 13 | E | K | R | R | N | K | R | Q | K | Q | K | Q | R | N | E | E | E | n.d. |
| 14 | Q | K | E | Q | N | Q | K | R | R | E | Q | Q | R | V | K | E | K | Q |

^a Sequence of library A: VAQL(e1)E(g1) VKTL(e2)A(g2) (a3)YEL(e3)S(g3) VQRL(e4)E(g4) VAQL.

^b Sequence of library B: VDEL(e'1)A(g'1) VDQL(e'2)D(g'2) (a'3)YAL(e'3)T(g'3) VAQL(e'4)K(g'4) VEKL.

hydrogen (Q, E, respectively), and the N3 side-chain should form a hydrogen bond with the N-cap amide hydrogen. The C-cap has only a minor effect on helix stability (Chakrabartty *et al.*, 1993). As Gly has a high preference for the C-cap position (Richardson & Richardson, 1988; Dasgupta & Bell, 1993), we added a C-terminal glycine. This may contribute to helix termination or extend the linker which connects the coiled coil to the DHFR fragments (Figure 1(b)).

Library screening strategy

Clones resulting from the three selection strategies (Pelletier *et al.*, 1999) with increasing stringencies (see Materials and Methods) were analyzed and compared: (i) lowest stringency, 14 clones analyzed (Table 1); (ii) medium stringency, 25 clones analyzed (Table 2); (iii) highest stringency, 41 clones analyzed from various passages (Table 3). The last passage (P12) yielded a population domi-

Table 2. Sequence pairs selected in the medium stringency selection (single-step selection using DHFR-I114A)

| Clone | Library A ^a | | | | | | | | | Library B ^b | | | | | | | | |
|-------|------------------------|----|----|----|----|----|----|----|----|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| | e1 | g1 | e2 | g2 | a3 | e3 | g3 | e4 | g4 | e'1 | g'1 | e'2 | g'2 | a'3 | e'3 | g'3 | e'4 | g'4 |
| 1 | Q | K | Q | E | N | E | R | R | K | K | E | E | R | N | K | E | R | Q |
| 2 | E | K | E | E | N | K | R | R | E | R | E | Q | Q | N | K | E | E | Q |
| 3 | K | Q | E | R | N | R | E | R | K | K | Q | E | E | N | E | Q | E | Q |
| 4 | R | K | Q | R | N | K | K | K | Q | Q | Q | R | E | N | E | Q | Q | R |
| 5 | Q | Q | K | R | N | K | R | Q | Q | Q | Q | E | Q | N | Q | Q | R | E |
| 6 | Q | R | K | E | N | K | R | R | K | K | E | Q | E | N | E | R | Q | R |
| 7 | E | E | E | R | N | R | K | E | Q | K | R | Q | K | N | R | K | Q | R |
| 8 | K | E | E | Q | N | K | K | K | K | K | K | Q | E | N | R | K | K | K |
| 9 | K | E | K | Q | N | R | E | E | Q | Q | E | Q | E | N | E | R | K | E |
| 10 | E | Q | Q | Q | N | K | Q | R | K | R | K | R | R | N | K | E | Q | E |
| 11 | K | R | R | E | N | R | E | E | R | K | K | R | E | N | K | R | Q | E |
| 12 | E | K | K | Q | N | E | E | R | R | Q | Q | Q | K | N | K | R | R | E |
| 13 | E | K | R | Q | N | E | Q | R | K | E | R | K | E | N | Q | E | E | E |
| 14 | R | R | K | E | N | E | R | K | E | R | R | E | K | N | K | Q | Q | R |
| 15 | R | R | Q | R | N | Q | Q | E | Q | K | K | R | E | N | E | Q | R | R |
| 16 | Q | E | R | E | N | R | K | K | Q | Q | R | R | E | N | R | K | Q | K |
| 17 | Q | K | E | K | N | E | E | E | R | K | K | K | R | N | K | Q | E | R |
| 18 | Q | R | Q | K | N | Q | Q | E | E | Q | K | E | Q | N | K | E | Q | E |
| 19 | Q | K | Q | K | N | K | Q | E | R | K | E | K | K | N | K | K | Q | R |
| 20 | Q | E | K | E | N | E | R | R | K | K | Q | Q | Q | N | R | Q | R | R |
| 21 | K | R | E | E | N | R | Q | E | R | K | R | E | E | N | R | R | E | R |
| 22 | K | E | E | R | N | R | R | E | R | K | Q | R | E | N | E | E | K | Q |
| 23 | K | Q | E | R | N | K | K | E | Q | R | R | Q | K | N | R | R | Q | Q |
| 24 | K | R | Q | Q | V | E | E | E | K | R | Q | E | K | N | Q | K | R | E |
| 25 | Q | K | R | R | V | E | Q | K | E | E | Q | E | Q | T | Q | E | E | E |

^a Sequence of library A: VAQL(e1)E(g1) VKTL(e2)A(g2) (a3)YEL(e3)S(g3) VQRL(e4)E(g4) VAQL.

^b Sequence of library B: VDEL(e'1)A(g'1) VDQL(e'2)D(g'2) (a'3)YAL(e'3)T(g'3) VAQL(e'4)K(g'4) VEKL.

Table 3. Sequence pairs selected in the highest stringency selection (competition selection)

| Clone | Clone frequency ^a | | Library A ^b | | | | | | | | Library B ^c | | | | | | | | | |
|-------|------------------------------|-----|------------------------|----|----|----|----|----|----|----|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | P12 | P10 | e1 | g1 | e2 | g2 | a3 | e3 | g3 | e4 | g4 | e'1 | g'1 | e'2 | g'2 | a'3 | e'3 | g'3 | e'4 | g'4 |
| 1 | 18 | 4 | E | K | R | Q | N | K | R | R | Q | Q | E | Q | E | N | K | K | R | K |
| 2 | 1 | 1 | Q | K | R | R | N | K | E | E | K | E | E | K | Q | N | K | K | Q | Q |
| 3 | 1 | 1 | E | Q | Q | R | N | K | K | K | K | R | K | Q | E | N | E | E | Q | R |
| 4 | 1 | | E | R | R | Q | N | K | K | E | Q | E | E | E | Q | N | Q | R | E | R |
| 5 | 1 | | E | Q | E | E | N | K | K | R | R | K | K | K | K | N | R | K | R | K |
| 6 | | 1 | Q | Q | E | Q | N | E | E | K | Q | R | Q | Q | K | N | R | R | K | R |
| 7 | | 1 | E | R | K | E | N | E | E | K | R | Q | E | Q | Q | N | R | Q | K | K |
| 8 | | 1 | E | K | K | K | N | K | K | K | R | Q | Q | E | Q | N | E | Q | E | E |
| 9 | | 1 | Q | E | Q | E | N | R | E | E | E | Q | K | K | E | N | Q | K | Q | R |
| 10 | | 1 | R | R | R | R | N | Q | K | K | R | R | E | E | E | N | R | R | R | Q |
| 11 | | | E | I | Q | E | N | E | Q | Q | K | E | E | E | R | N | R | Q | K | R |
| 12 | | | E | Q | Q | E | N | Q | Q | Q | E | Q | R | Q | K | N | K | E | K | R |
| 13 | | | K | K | R | E | N | Q | E | R | R | K | Q | Q | E | N | E | E | R | K |

^a The frequency and distribution of each clone in passages 12 and 10 is shown. Clone 6 also was found twice in P8; clones 2 and 7 were each found once in P6; clone 11 was found twice in P4; clones 12 and 13 were each found once in P2.

^b Sequence of library A: VAQL(e1)E(g1) VKTL(e2)A(g2) (a3)YEL(e3)S(g3) VQRL(e4)E(g4) VAQL.

^c Sequence of library B: VDEL(e'1)A(g'1) VDQL(e'2)D(g'2) (a'3)YAL(e'3)T(g'3) VAQL(e'4)K(g'4) VEKL.

nated by a single pair of coiled-coil sequences, WinZip-A1B1 (Figure 1(c), clone no. 1 in Table 3), which was biophysically analyzed (see below). By comparing the selected clones from the three strategies, we analyzed the preferences for the core a position (Pelletier *et al.*, 1999), the distribution of e/g pair combinations and the presence of any bias for certain amino acids within the individual helices as detailed below.

Selection for stability and specificity

All selected clones must form a heterodimer to reconstitute DHFR activity. However, the heterospecificity, and thus the ratio of heterodimers (library A/library B) to homodimers (library A/library A and library B/library B) can vary. The homodimers cannot reconstitute active enzyme, and may "titrate out" fragments from functional assembly. Accumulation of non-functional protein might even be harmful. However, a mixture of homo- and heterodimers may generate sufficient amounts of active enzyme to survive under low stringency conditions, but not under competitive growth conditions (high stringency selection).

We analyzed the average occurrence of all e/g-pair combinations in the heterodimers and the putative homodimers arising from the various selections. In a parallel helix arrangement, six e/g-pair interactions ($g_1-e'_2$, $g_2-e'_3$, $g_3-e'_4$, $e_2-g'_1$, $e_3-g'_2$, $e_4-g'_3$) are possible in each dimer (Figure 1(a)). The e/g interactions can be potentially attractive (E-K, K-E, E-R, R-E), neutral (Q-Q, Q-E, E-Q, Q-K, K-Q, Q-R, R-Q) or repulsive (E-E, K-K, K-R, R-K, R-R). The average of each interaction type in the entire population arising from each selection is reported in relation to the random statistical distribution (25% attractive, 44% neutral, 31% repulsive) (Figure 2).

The selected heterodimers (Figure 2(a)) show on average more attractive and less repulsive inter-

actions than expected in a random population, indicating selection for stability. This trend, although increasing with higher stringency, is already observed in the lowest stringency selection, indicating that a certain threshold of stability is needed to induce enzyme complementation. Selection for heterospecificity is achieved by destabilizing the homodimers (Figure 2(b)) relative to the heterodimer (Figure 2(a)) and is mainly observed in the medium and highest stringency selections using the destabilizing I114A DHFR[2]-mutant. Interestingly, this effect is more pronounced for library A homodimers than for library B homodimers (data not shown), and the biophysical characterization of WinZip-A1B1 indicated that the library A homodimer is more stable (see below), and thus might have a stronger influence on titrating out fragments than the library B homodimer. However, it is worth noting that the predominantly selected pair, WinZip-A1B1, contains two repulsive e/g interactions. Thus, while an enrichment of attractive interactions is seen in the library-versus-library population, the complete absence of repulsive interactions is obviously not a requirement for the most efficient performance among the library pool members. Furthermore, in the medium as well as in the highest stringency selection, 13 out of 38 pairs sequenced had no repulsive e/g pairs, but none competed successfully against WinZip-A1B1 in the selection.

To determine the degree of heterospecificity achieved in the various selections, the relative stability of heterodimer *versus* homodimer was estimated for each single clone. We calculated the difference of attractive or repulsive interactions, respectively, between the heterodimer and the average of the two corresponding homodimers (Figure 3). The results clearly show that heterospecificity is achieved not only by a decrease of repulsive interactions but also by an increase of attractive interactions in the heterodimer relative to

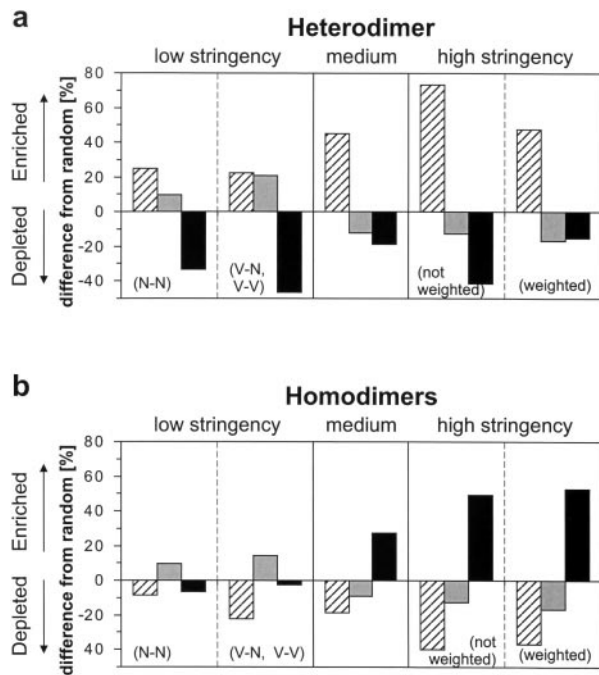


Figure 2. Deviation of observed *e/g* interactions in the selected heterodimer (a) and the two putative homodimers (b) from the statistically expected distribution in the absence of selection. Interactions are potentially attractive (E-K, K-E, E-R, R-E; hatched bars), neutral (Q-Q, Q-E, E-Q, Q-K, K-Q, Q-R, R-Q; grey bars) or repulsive (E-E, K-K, K-R, R-K, R-R; black bars). Low stringency selection: clones ($n = 14$, Table 1) were subdivided into those with an N-N pair in the core a position ($n = 8$) and those with an N-V or V-V pair ($n = 6$). Medium stringency selection: only clones with an Asn pair in the core a position were considered ($n = 23$, clones 1-23 in Table 2). High stringency selection: clones which survived the competition selection at least up to passage 10 (clones 1-10 in Table 3) were considered. These were analyzed counting each sequence once (not weighted, $n = 10$) or according to their frequency of appearance (weighted, $n = 37$). For example, the high stringency selection yielded in total 82 attractive (opposite charges) interactions in the 37 sequences analyzed. In a random distribution (37 pairs with six *e/g* interactions each), 25% or 55.5 *e/g* interactions would be expected to be attractive (opposite charges). Thus, the observed number of attractive pairs exceeds the statistically expected one by 48%.

the homodimers. However, the lowest (Figure 3(a)) and medium stringency selection (Figure 3(b)) yielded still a certain fraction of pairs with no difference or even a reversed ratio, whereas the highest stringency selection (Figure 3(c) and (d)) exclusively yielded pairs with distinct heterospecificity. In addition, in no case were more than three repulsions found in the heterodimers, although in a random combination 8% of all pairs should have four to six repulsions.

Thus, the data indicate that a certain threshold of stability is needed for successful enzyme complementation, and that heterospecificity is the most

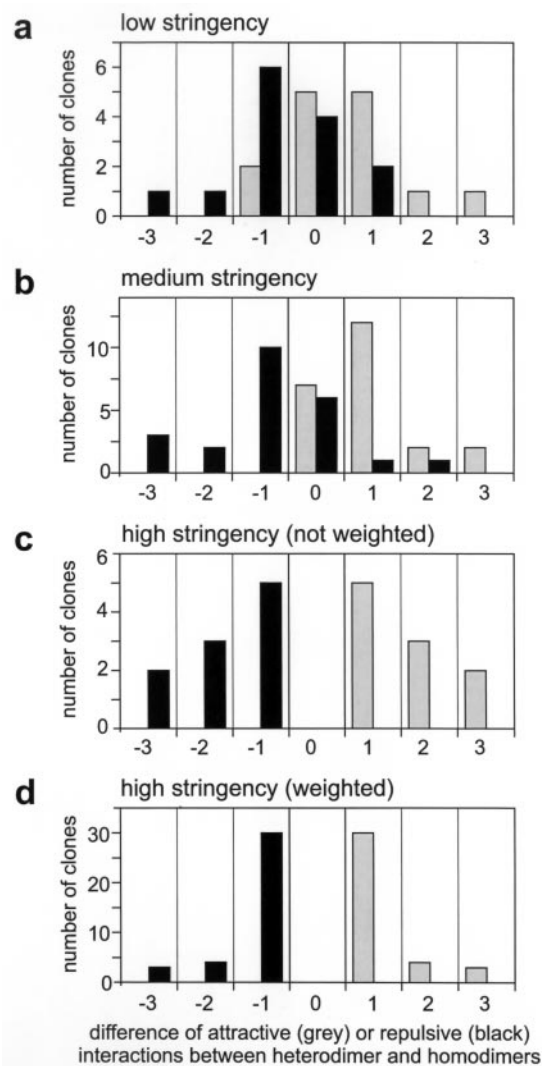


Figure 3. Number of selected pairs having a certain difference of attractive (grey bars) or repulsive interactions (black bars), respectively, between the heterodimer and its constituting homodimers. For example, a value of -3 for the repulsive interactions (black bars) indicates that there are three fewer potentially repulsive interactions in the heterodimer than in the average of both homodimers. (a) Fourteen clones from the low stringency selection, (b) 23 clones from the medium stringency selection, (c) ten different clones from the high stringency selection (not weighted), (d) 37 clones from the high stringency selection (weighted).

important selection criteria in the more stringent selections. The importance of specificity was also observed in the core a position (as described by Pelletier *et al.*, 1999). The strong selection for Asn pairs at this position (57% of 14 pairs analyzed from the low stringency selection, 92% of 25 pairs from the medium and 100% of 41 pairs from the highest stringency selection) indicated that the specificity gained by this polar interaction seems to be a very important feature, since it clearly out-

weighs the more energetically stable Val pairs even in the low stringency selection.

Positional distribution of the selected amino acids

Intrahelical electrostatic interactions can influence stability and may even promote selection of apparently repulsive e/g pairs. Interactions with the helix macrodipole, for example, can modulate stability in coiled coils (Kohn *et al.*, 1997). Indeed, we observed a bias for negatively charged and neutral amino acids in the N-terminal part and positively charged amino acids in the C-terminal part (Figure 4). This positional preference may at least partially compensate the loss of stability resulting from a repulsive e/g interaction. In addition, interactions with adjacent residues on the outside of the helix (b and c positions) may influence the contributions of charges at the e and g positions (Yu *et al.*, 1996). This may explain why at position e1 in library B (compare e1 in Figure 4(b) with e1 in Figure 4(a)) a negatively charged amino acid is not favored, contrary to the expected counterbalancing of the helix dipole, since this position is adjacent to two aspartate residues in positions b1 and b2. On a more general note, the predominantly selected sequence with the residues from c-Jun at the outer positions (WinZip-A1, selected from library A) bears remarkable similarity to the e and g sequences in the naturally occurring c-Jun (compare WinZip-A1: E-K; R-Q; K-R; R-Q, with c-Jun: E-K; K-Q; A-T; R-Q; order of residues shown, e1-g1; e2-g2; e3-g3; e4-g4).

Library complexity

Although the predominantly selected sequence pair WinZip-A1B1 showed all the desired properties *in vivo* as well as *in vitro* (see below), we were not able to cover all theoretical library-versus-library combinations in our selection. Nonetheless, we covered all possible electrostatic interaction combinations (+/+ , -/- , +/- , +/-n , -/n, n/n; n = neutral) in all six interacting e/g positions, when grouping the core a position into favored (Asn-Asn) or disfavored (Asn-Val, Val-Val) combinations. This reduces the theoretical library size of 1.7×10^{10} possibilities to 9×10^4 possibilities of different charge combinations relevant for the presented studies. This number was well covered by the experimental library of 2×10^6 . It is therefore likely that WinZip-A1B1 contains the most important features for stability and heterospecificity, although further optimization may be possible. Furthermore, the random probability of finding pairs with no repulsive interactions was 1:40, and with solely attractive interactions was $1:1.6 \times 10^4$. Thus, our selection covered a representative sequence space and the same-charged interactions in WinZip-A1B1 are not a result of incomplete library sampling but must result from more subtle reasons, including *in vivo* factors that

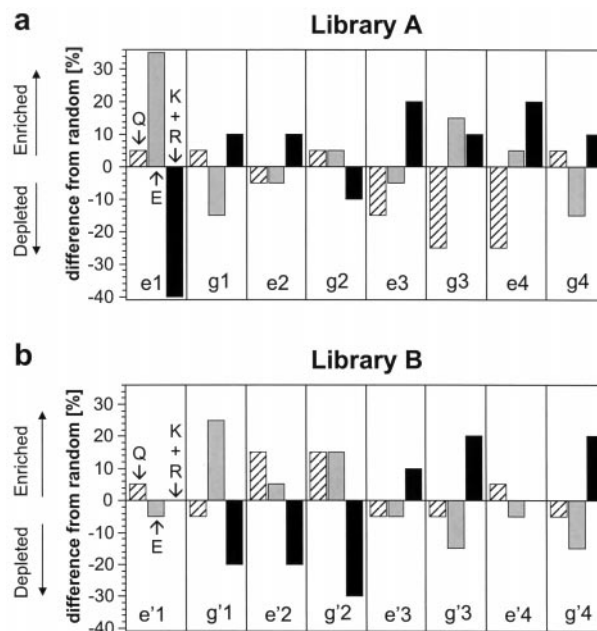


Figure 4. Positional distribution of amino acids at each e and g position in sequences from library A (a) and library B (b), obtained from the highest stringency selection. The statistically expected random occurrence of each amino acid at each position was subtracted from the relative occurrence observed in the selection (Q is hatched, E is grey, K/R is black).

we cannot fully address with theoretical considerations.

Secondary structure and oligomerization state of the predominant pair WinZip-A1B1

We investigated the stability and specificity of the predominantly selected peptides WinZip-A1 and WinZip-B1 alone and in an equimolar mixture (WinZip-A1B1). All experiments were performed with N-acetylated and C-amidated synthetic peptides. All peptides formed stable α -helical coiled coils as demonstrated by CD-spectra (Figure 5(a)). The helical content was in the range of 90% (WinZip-B1) to 100% (WinZip-A1 and WinZip-A1B1). Peptide WinZip-A1 (data not shown) as well as the mixture WinZip-A1B1 (Figure 6) were dimeric at 10°C and 25°C over a concentration range from 10 to 150 μ M as determined by equilibrium sedimentation. WinZip-B1 was partially unfolded as seen both by CD (Figure 5(a)) and equilibrium sedimentation (data not shown), which indicated a mixture of monomers and dimers, with decreasing amount of dimer at increasing temperature.

Structural stability and heterospecificity

Thermal denaturation studies at neutral pH and 150 μ M total peptide concentration (Figure 5(b)) revealed apparent t_m values of 28°C (WinZip-B1), 49°C (WinZip-A1), and 55°C for the equimolar

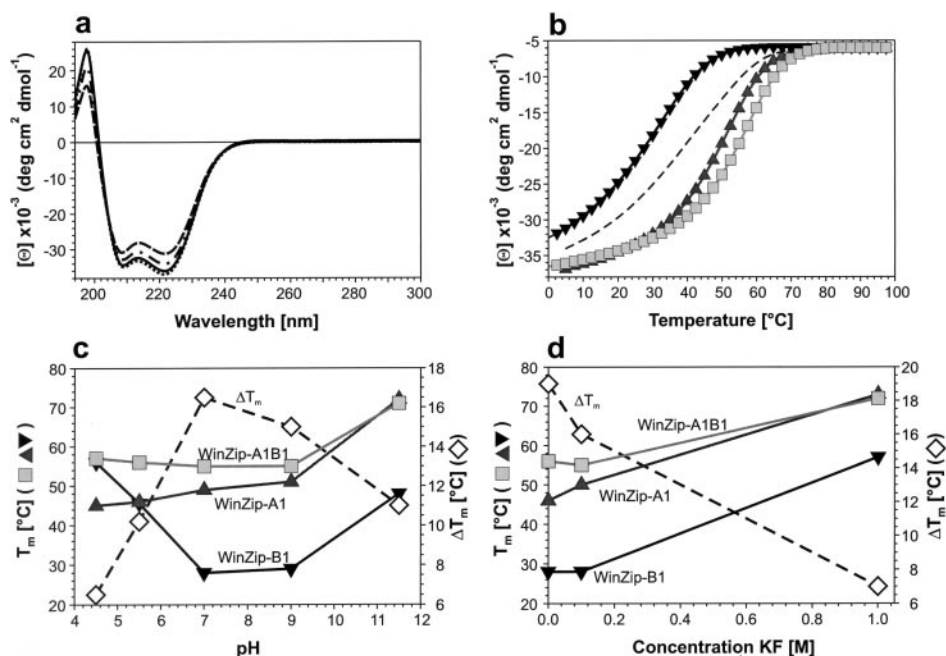


Figure 5. CD measurements of the synthesized peptides of WinZip-A1B1. (a) Far-UV CD-spectra of 150 μM total peptide concentration at 5°C for the heterodimer WinZip-A1B1 (—), the homodimers WinZip-A1 (·····) and WinZip-B1 (---), and the calculated average between WinZip-A1 and WinZip-B1, which would be expected assuming no interaction (-·-·-). (b) Temperature dependence of $[\theta]_{222}$ for WinZip-A1B1 (\square , light grey), WinZip-A1 (\blacktriangle , dark grey), WinZip-B1 (\blacktriangledown , black), and the calculated average of both homodimers (---). (c) Dependence of t_m and Δt_m (\diamond , white) on pH, and (d) on salt. The t_m was calculated from thermal denaturation curves (same symbols are used as in panel (b)). The Δt_m was calculated as the difference between the t_m of the heterodimer and the averaged t_m of both homodimers. It must be noted that thermal denaturation of WinZip-A1 was not completely reversible at 1 M salt concentration, and only 71% of the starting signal was regained.

mixture of both (WinZip-A1B1). The large difference between the denaturation curve of the heterodimer and the average of the curves from WinZip-A1 and WinZip-B1 indicates that heterodimers form preferentially at equilibrium. This high heterospecificity is best reflected in a large and positive Δt_m value, and indeed, we observed a Δt_m of 16.5 deg. C.

To probe the mechanism of specificity, the effects of pH (Figure 5(c)) and ionic strength (Figure 5(d)) were investigated. All peptides were more stable at high pH, most likely because all have at least one e/g pair with two positive charges which are neutralized at high pH. The increased stability of WinZip-B1 at low pH could be due to the shielding of electrostatic repulsions resulting from its high concentration of acidic residues. However, the Δt_m is positive over the whole pH range, indicating heterospecificity. The maximum degree of heterospecificity was observed at neutral to slightly basic pH, consistent with the intracellular pH of *E. coli*, which is between pH 7.5 and 7.9 (Oliver, 1996). High salt concentrations increased the absolute t_m values (Figure 5(d)), presumably due to increased hydrophobic interaction and reduced electrostatic repulsion. However, the Δt_m is reduced compared to low salt concentrations (0–100 mM), most likely due to the decreased influence of ionic interactions at higher ionic strength.

Interestingly, the overall stability did not correlate directly with the number of potentially repulsive e/g interactions. The homodimer WinZip-B1 has two same-charged ion pairs, but is significantly less stable than the homodimer WinZip-A1 with four same-charged pairs (Figures 5(b) and 8). Since the overall helical propensity is comparable for both peptides, estimated according to the scale described by O'Neil & DeGrado (1990), the stability difference is probably due to intrahelical interactions. Library B might be destabilized by its high local concentration of acidic residues at the N terminus. This may also explain why library A homodimers from the selected heterodimer pairs have generally more repulsive and less attractive e/g interactions than library B homodimers, since the e/g positions play a more important role in the destabilization of the intrinsically more stable library A in order to reduce homodimerization.

Heterospecificity also was observed by native gel electrophoresis (Figure 7). To obtain a significant migration, an acidic buffer (pH 4.5) had to be used. Under these conditions where the heterospecificity is the lowest (Δt_m of only 7 deg. C, compared to 16.5 deg. C at neutral pH, Figure 5(c)). Nevertheless, even under these stringent conditions, heterodimers were obtained almost exclusively from the equimolar mixture (Figure 7), suggesting very high heterospecificity at neutral

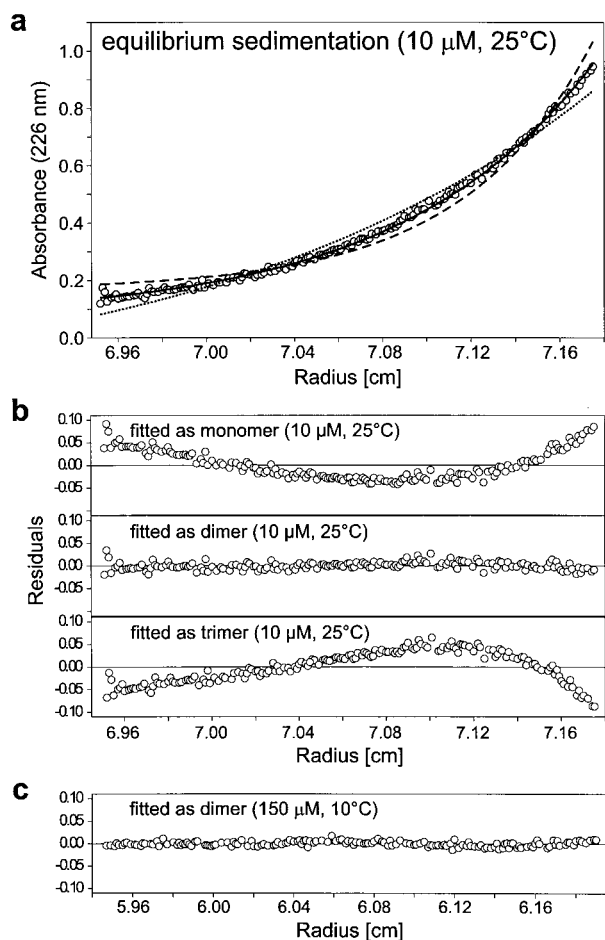


Figure 6. Determination of the molecular mass of WinZip-A1B1 by sedimentation equilibrium. (a) Measured data (○) obtained at 10 μM peptide concentration at 25°C, as well as the fit for a monomer (\cdots), dimer (—), or trimer (---). (b) Residuals between measured data obtained at 10 μM peptide concentration at 25°C and data fitted as monomer (top), dimer (middle), or trimer (bottom). (c) Residuals for a dimeric fit of the data set for WinZip-A1B1 obtained at 150 μM peptide concentration, 10°C, demonstrating that over this concentration and temperature range, the WinZip-A1B1 is dimeric.

pH, and thus indicating how strongly heterospecificity was selected.

K_D determination

Dissociation constants of the peptides were derived from equilibrium urea denaturations (Figure 8). The heterodimer WinZip-A1B1 was the most stable species, with a K_D of approximately 24 nM. The homodimer WinZip-A1 had a K_D of approximately 63 nM. The accuracy of the K_D determination of WinZip-B1 is lower since it is partially unfolded without denaturant (see above). The K_D of WinZip-B1 was estimated to be in the 10^{-5} M range. Calculations were confirmed by determining the K_D values from thermal denatura-

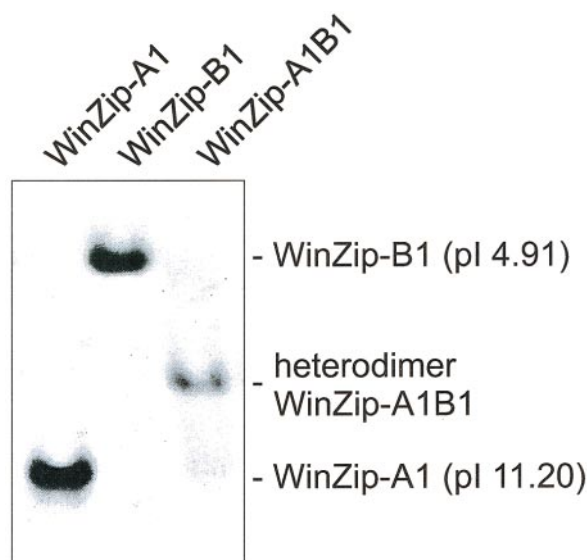


Figure 7. Native gel electrophoresis of WinZip-A1, WinZip-B1 and WinZip-A1B1, performed as described in Materials and Methods.

tion curves by a van't Hoff analysis, assuming as a first approximation a constant ΔH (Becktel & Schellman, 1987). We found reasonable agreement to the data obtained by urea denaturation with a maximal deviation of K_D by a factor of 2.6 (WinZip-A1B1, ~ 64 nM; WinZip-A1, 137 nM; WinZip-B1, 10^{-5} M range).

Comparison to other coiled coils

Designed coiled coils usually are judged only for being stable *in vitro* and, in certain cases, for heterospecificity, whereas naturally occurring coiled coils also must function reliably in a cellular environment. Similar demands are imposed during our

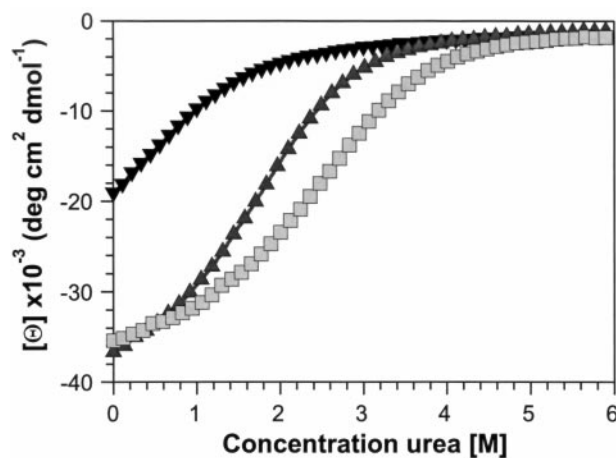


Figure 8. Urea titration of the heterodimer WinZip-A1B1 (□, light grey), and the homodimers WinZip-A1 (▲, dark grey) and WinZip-B1 (▼, black).

selection and in further *in vivo* applications, and therefore WinZip-A1B1 is compared best with peptides from other naturally occurring coiled coils. The homodimeric coiled coil of the yeast transcription factor GCN4 has an equal to slightly higher t_m , depending on the length and concentration of the peptides chosen (Thompson *et al.*, 1993; Lumb *et al.*, 1994). The N-terminal, homodimeric coiled coil of the APC protein has a t_m lower by at least 9 deg. C than WinZip-A1B1, even though the characterized peptide of APC was longer and more highly concentrated (Joslyn *et al.*, 1993). The coiled coil from c-Myc/Max heterodimerizes to a fairly high extent, but peptides of comparable length have a t_m of only 31 °C and a K_D of 60 μ M (25 °C) (Muhle-Goll *et al.*, 1995), whereas our WinZip-A1B1 has a t_m of 55 °C and a K_D of 24 nM (20 °C). The heterodimeric coiled coil from c-Jun/c-Fos shows comparable t_m and Δt_m values to WinZip-A1B1 (O'Shea *et al.*, 1992). However, those data were derived from disulfide-bridged peptides, and thus represent an intramolecular reaction, whereas we observed a true intermolecular reaction. Thus, WinZip-A1B1 compares favorably with naturally occurring coiled coils and will therefore be very useful for a variety of *in vivo* applications.

Discussion

We applied and analyzed a fast and simple strategy, a library-versus-library selection with the fragment complementation assay, to select for a metabolically stable, dimeric and highly heterospecific coiled coil with high affinity. Comparison of the outcome of various selections performed with different stringencies revealed insight into which properties are selected already at lower stringency, and are thus the most crucial for successful heterodimerization, and those which only become apparent at higher selection stringency, and thus represent a more subtle optimization. The most striking selection occurred at the core a position for Asn pairs, revealing that structural uniqueness is essential for efficient and selective heterodimerization. Furthermore, comparison of selected e/g pairs from heterodimers and putative homodimers indicated selection for stability even at the lowest stringency, whereas selection for heterospecificity was more pronounced at higher stringency. Heterospecificity was achieved not only by decreasing the numbers of repulsive e/g interactions but also by increasing the number of attractive interactions in the heterodimer relative to the homodimers.

The selection for heterospecificity (and thus against homodimers) may be a unique feature of this selection system. Not only is active enzyme exclusively formed by parallel heterodimers, but homodimers and higher-order oligomers are likely to have a negative effect by unproductively wasting fragments, and perhaps accumulation of non-functional enzyme may even be harmful. Dimer stability, in turn, depends not only on e/g pair

interactions, but also on helical propensity, intrahelical interactions and helix dipole stabilization. Indeed, our analysis revealed that the most successful variants do not simply contain complementary charges in the e/g positions, but show a more complicated pattern, presumably fulfilling a variety of naturally conflicting demands on the sequence. This optimum would have been extremely challenging to predict.

The biophysical characterization of the predominantly selected pair WinZip-A1B1 revealed the formation of a stable, dimeric coiled coil with very high heterospecificity. These results confirmed the sequence analysis and the validity of the selection strategy. The heterodimer was found to be the most stable species (Figure 5) and formed almost exclusively from an equimolar mixture of both homodimers (Figure 7). The homodimer WinZip-B1 is relatively unstable, whereas the homodimer WinZip-A1 is almost, but not quite, as stable as the heterodimer. This behavior is very similar to Jun/Fos-heterodimerization (O'Shea *et al.*, 1989). Currently, we are analyzing the biophysical results of chain shuffling experiments, in which either one of the helices of WinZip-A1B1 was kept constant and selected against the entire library of complementary helices. According to preliminary results the observed e/g interaction pattern is similar to that of WinZip-A1B1, supporting the view that idealized sequences, based on the single principle of merely relieving repulsive e/g interactions in the homodimers with complementary charges in the heterodimer, may not be optimal for biological applications.

In conclusion, the identification of the heterodimeric coiled coil WinZip-A1B1 demonstrated that the combination of designed well-understood features in the helices, together with the combinatorial "fine tuning" of more subtle interactions along and across the helices, is a very powerful technology, and is probably the most fruitful strategy for the design of heterodimeric peptides currently available.

Materials and Methods

Library synthesis

Trinucleotide codons (Virnekäs *et al.*, 1994) were used to code for randomized positions; all other positions were made with mononucleotides. Library A (104 bp), 5'-TACTGTGGCGCAACTGNNNGAANNNGTGAAAA-CCCTTNNNGCTNNNXXXATGAACTTNNNTCT-NNNGTGCAGCGCTTGNNNGAGNNNGTTGCCAG-CTTGCTA-3'; library B (106 bp), 5'-CTCCGTTGAC-GAACTGNNNGCTNNNGTTGACCAGCTGNNNGAC-NNNXXXTACGCTCTGNNNACCNNNGTTGCGCAGC-TGNNNAAAANNNGTGAAAAGCTGTGATAA-3' (NNN = equimolar mixture of the trinucleotides AAG (Lys), CAG (Gln), GAG (Glu), CGT (Arg); XXX = equimolar mixture of the trinucleotides AAT (Asn), GTT (Val)). Generation of the second strand and introduction of *SalI* and *NheI* restriction sites were achieved by PCR using the primers prA-fwd, 5'-GGAGTACTGGCATGCAGTC-

GACTACTGTGGCGCAACTG-3', and prA-rev, 5'-GGACTAGTACCTTCGCTAGCAAGCTGGGCAAC-3', for library A and the primers prB-fwd, 5'-GGAGT-ACTGGCATGCAGTCGACCTCCGTTGACGAACTG-3', and prB-rev, 5'-GGACTAGTGCTAGCTTCTGACAGCTTTTCCAC-3', for library B. This resulted in a 142 bp double-stranded oligonucleotide for either library.

Expression plasmids

Library A and library B were both cloned *via* *Sall*/*NheI* into a variant of plasmids Z-F[1,2], encoding the N-terminal DHFR fragment, and Z-F[3] or Z-F[3:I114A], respectively, encoding the C-terminal DHFR fragment with or without the I114A mutation (Pelletier *et al.*, 1998), yielding the plasmids LibA-DHFR[1], LibB-DHFR[2], LibB-DHFR[2:I114A] as described by Pelletier *et al.* (1999). In both plasmids, an N-terminal His₆-tag (underlined) followed by a linker sequence (MRGSHHHHHHGIRMRAR-YPGSTT) precedes each library. A second linker (ASGTSSGTSSTSSGI for library A and SEASGTSSGTSSTS for library B) connects the libraries with the respective DHFR fragment. Libraries were electroporated separately into XL1-blue cells yielding about 10⁶ clones per library. Cells from each library were pooled, and their plasmid DNA was cotransformed in BL21 cells yielding about 4 × 10⁶ double-transformants. We observed a cotransformation rate of approximately 50% when using 0.5 μg of DNA for each plasmid (Pelletier *et al.*, 1999). Sequencing of individual clones prior to selection revealed an equal distribution of amino acids at the randomized positions, and 70-80% of each library was without frameshifts. Thus, the experimental library-versus-library size of correct pairs is estimated as 2 × 10⁶.

Selection procedure

The lowest and medium stringency selections used the single-step protein complementation assays as described by Pelletier *et al.* (1998), where cells cotransformed with complementary libraries were directly plated on selective medium plates (M9 medium with 1 μg/ml trimethoprim), and resulting colonies were analyzed. The lowest stringency selection used the wild-type DHFR fragments, the medium stringency selection used the destabilizing I114A mutation in DHFR[2] which occurs at the interface between both DHFR fragments. In the highest stringency selection (competition selection; Pelletier *et al.*, 1999), the clones resulting from the medium stringency selection were pooled and cultured in liquid medium under selective conditions over 12 passages (serial transfers). The most stable heterodimers should result in higher mDHFR activity and thus provide a growth advantage. DNA sequences were obtained by automated sequencing.

Peptide synthesis and purification

The peptides WinZip-A1, Ac-STTVAQLEEK-VKTLRAQNYELKSRVQRLREQVAQLAS-NH₂, and WinZip-B1, Ac-STSVDLQAEVDQLQDENYALKTKVAQLRKKVEKLSE-NH₂, were synthesized using solid-phase methodology with an Applied Biosystem 431A automated synthesizer. Both peptides were acetylated at their N terminus and amidated at their C terminus in order to closer resemble the conditions in the DHFR fusions. In addition, to allow helix capping and increase

solubility rather than starting and ending with a hydrophobic amino acid, three N and two C-terminal residues were added. Those amino acids are identical with those preceding and succeeding the coiled coil sequence in the fusion protein. Peptide resin cleavage and peptide deprotection were accomplished in a single step using reagent K (King *et al.*, 1990) for four hours at room temperature. The crude peptide material was purified by reversed-phase HPLC on a semipreparative (1.0 cm × 25 cm) C18 column (Vydac, CA) using linear gradients of acetonitrile in water, both containing 0.1% TFA. Identity and purity of the peptides was confirmed by electrospray mass spectrometry. The observed mass deviation was less than 1 Da for both peptides (data not shown). Peptide concentrations were determined by tyrosine absorbance in 6 M GdnHCl (Edelhoch, 1967).

Circular dichroism measurements

CD studies were performed with an Aviv model 62DS spectrometer. Spectra of each peptide alone as well as of the respective equimolar mixture were measured at 5°C using a total peptide concentration of 150 μM in a 1 mm cuvette. The standard buffer was 10 mM potassium phosphate (pH 7.0), 100 mM KF; salt concentration and pH were varied as indicated for the respective experiments. Thermal denaturations were measured at 222 nm from 0°C to 97.5°C in steps of 2.5 deg. C (two minutes equilibration, 30 seconds data averaging). Thermal transitions were >91% reversible except where indicated. Apparent *t_m* were determined by least-squares curve fitting of the denaturation curves (Becktel & Schellman, 1987), assuming a two-state model (folded dimer, unfolded monomer). Δ*t_m* was calculated as *t_m* (WinZip-A1B1) - 1/2 [*t_m*(WinZip-A1) + *t_m*(WinZip-B1)]. Urea denaturation equilibria were determined at 20°C by automated titration of native peptide with denatured peptide in 6 M urea (30 μM WinZip-A1 or WinZip-A1B1, respectively, or 60 μM WinZip-B1) measuring the CD signal at 222 nm (300 seconds equilibration, 30 seconds data averaging). *K_D* values were calculated by linear extrapolation to 0 M denaturant assuming a two-state model (*K_D* = [unfolded monomer]²/[folded dimer]).

Equilibrium sedimentation

Equilibrium sedimentation experiments were performed using a Beckman XL-A Ultracentrifuge equipped with absorbance optics and a Beckman AN-50 rotor. Samples were extensively dialyzed against 10 mM potassium phosphate buffer (pH 7.0), 100 mM KCl and loaded at three concentrations (10, 50, 150 μM) into a six-hole centerpiece and centrifuged at 40,000 rpm for at least 24 hours at 10°C and 20°C. Five to seven data sets were obtained at 220 nm and 275 nm. Partial specific volumes and solvent densities were determined as described (Laue *et al.*, 1992). The data sets were fitted to single molecular masses of monomer, dimer and trimer.

Native gel electrophoresis

Gels (7.5% (w/v) polyacrylamide (19:1 (w/w) acrylamide/bis-acrylamide), 375 mM β-alanine acetate buffer, pH 4.5) were run with 500 mM β-alanine acetate buffer (pH 4.5). Samples (~10 μg peptides per lane) were two-fold diluted with 600 mM β-alanine acetate (pH 4.5), 0.2% (w/v) Methyl green, 30% (v/v) glycerol. Gels were

prerun at 100 V for at least 45 minutes and run for two to three hours at 5 °C. The electrodes were reversed compared to SDS gel electrophoresis, since all peptides were positively charged under the conditions chosen. Gels were fixed with 2% (w/v) glutaraldehyde or 20% (w/v) TCA, respectively, before staining with Coomassie blue.

Acknowledgments

This work was supported by grant no. 0311628 from the German Bundesministerium für Bildung und Forschung (A.P.) and the Natural Sciences and Engineering Research Council of Canada (NSERC) (S.W.M.) and NIH grant GM48958 (to T.A.). K.M.A. is a recipient of a PhD grant from the Roche Research Foundation (formerly Stipendienfond der Basler Chemischen Industrie). J.N.P. was a recipient of a Fellowship from les Fonds de la Recherche en Santé du Québec, and currently holds a Fellowship from NSERC (Canada). K.M.M. holds a post-doctoral fellowship from the Schweizerische National Fund. S.W.M. is supported by a Burroughs Wellcome Fund New Investigator Award. We thank Günter Wellnhofer (Morphosys AG, Munich, Germany) for the synthesis of the oligonucleotides with trinucleotides, D. King for peptide synthesis and Michel Denault for his help in statistical analysis.

References

- Becktel, W. J. & Schellman, J. A. (1987). Protein stability curves. *Biopolymers*, **26**, 1859-1877.
- Chakrabarty, A., Doig, A. J. & Baldwin, R. L. (1993). Helix capping propensities in peptides parallel those in proteins. *Proc. Natl Acad. Sci. USA*, **90**, 11332-11336.
- Cohen, C. & Parry, D. A. (1990). Alpha-helical coiled coils and bundles: how to design an alpha-helical protein. *Proteins: Struct. Funct. Genet.* **7**, 1-15.
- Curran, T. & Franza, B. R., Jr (1988). Fos and Jun: the AP-1 connection. *Cell*, **55**, 395-397.
- Dasgupta, S. & Bell, J. A. (1993). Design of helix ends. Amino acid preferences, hydrogen bonding and electrostatic interactions. *Int. J. Pept. Protein Res.* **41**, 499-511.
- Doig, A. J. & Baldwin, R. L. (1995). N- and C-capping preferences for all 20 amino acids in alpha-helical peptides. *Protein Sci.* **4**, 1325-1336.
- Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948-1954.
- Graddis, T. J., Myszka, D. G. & Chaiken, I. M. (1993). Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry*, **32**, 12664-12671.
- Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, **262**, 1401-1407.
- Harbury, P. B., Kim, P. S. & Alber, T. (1994). Crystal structure of an isoleucine-zipper trimer. *Nature*, **371**, 80-83.
- Hodges, R. S. (1996). Boehringer Mannheim award lecture 1995. De novo design of alpha-helical proteins: basic research to medical applications. *Biochem. Cell Biol.* **74**, 133-154.
- Hu, J. C., Newell, N. E., Tidor, B. & Sauer, R. T. (1993). Probing the roles of residues at the e and g positions of the GCN4 leucine zipper by combinatorial mutagenesis. *Protein Sci.* **2**, 1072-1084.
- Hurst, H. C. (1995). Transcription factors 1: bZIP proteins. *Protein Profile*, **2**, 101-168.
- Jelesarov, I. & Bosshard, H. R. (1996). Thermodynamic characterization of the coupled folding and association of heterodimeric coiled coils (leucine zippers). *J. Mol. Biol.* **263**, 344-358.
- John, M., Briand, J. P., Granger-Schnarr, M. & Schnarr, M. (1994). Two pairs of oppositely charged amino acids from Jun and Fos confer heterodimerization to GCN4 leucine zipper. *J. Biol. Chem.* **269**, 16247-16253.
- Joslyn, G., Richardson, D. S., White, R. & Alber, T. (1993). Dimer formation by an N-terminal coiled coil in the APC protein. *Proc. Natl Acad. Sci. USA*, **90**, 11109-11113.
- Junius, F. K., Mackay, J. P., Bubb, W. A., Jensen, S. A., Weiss, A. S. & King, G. F. (1995). Nuclear magnetic resonance characterization of the Jun leucine zipper domain: unusual properties of coiled-coil interfacial polar residues. *Biochemistry*, **34**, 6164-6174.
- King, D. S., Fields, C. G. & Fields, G. B. (1990). A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int. J. Pept. Protein Res.* **36**, 255-266.
- Kohn, W. D., Kay, C. M. & Hodges, R. S. (1995). Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. *Protein Sci.* **4**, 237-250.
- Kohn, W. D., Kay, C. M. & Hodges, R. S. (1997). Positional dependence of the effects of negatively charged Glu side-chains on the stability of two-stranded alpha-helical coiled-coils. *J. Pept. Sci.* **3**, 209-223.
- Krylov, D., Mikhailenko, I. & Vinson, C. (1994). A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J.* **13**, 2849-2861.
- Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. M. (1992). In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90-125, Royal Society of Chemistry, Cambridge.
- Lumb, K. J. & Kim, P. S. (1995). A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry*, **34**, 8642-8648.
- Lumb, K. J., Carr, C. M. & Kim, P. S. (1994). Subdomain folding of the coiled coil leucine zipper from the bZIP transcriptional activator GCN4. *Biochemistry*, **33**, 7361-7367.
- Muhle-Goll, C., Nilges, M. & Pastore, A. (1995). The leucine zippers of the HLH-LZ proteins Max and c-Myc preferentially form heterodimers. *Biochemistry*, **34**, 13554-13564.
- Müller, K. M., Arndt, K. M. & Plückthun, A. (1998a). A dimeric bispecific miniantibody combines two specificities with avidity. *FEBS Letters*, **432**, 45-49.
- Müller, K. M., Arndt, K. M., Strittmatter, W. & Plückthun, A. (1998b). The first constant domain (C_H1 and C_L) of an antibody used as heterodimerization domain for bispecific miniantibodies. *FEBS Letters*, **422**, 259-264.
- Nautiyal, S., Woolfson, D. N., King, D. S. & Alber, T. (1995). A designed heterotrimeric coiled coil. *Biochemistry*, **34**, 11645-11651.

- Ogihara, N. L., Weiss, M. S., DeGrado, W. F. & Eisenberg, D. (1997). The crystal structure of the designed trimeric coiled coil VaLd: implications for engineering crystals and supramolecular assemblies. *Protein Sci.* **6**, 80-88.
- Oliver, D. B. (1996). Periplasm. In *Escherichia coli and Salmonella* (Neidhardt, F. C., ed.), vol. 1, pp. 88-103, ASM Press, Washington, DC.
- O'Neil, K. T. & DeGrado, W. F. (1990). A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science*, **250**, 646-651.
- O'Shea, E. K., Rutkowski, R., Stafford, W. F. & Kim, P. S. (1989). Preferential heterodimer formation by isolated leucine zippers from Fos and Jun. *Science*, **245**, 646-648.
- O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science*, **254**, 539-544.
- O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell*, **68**, 699-708.
- O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide "Velcro": design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658-667.
- Pack, P. & Plückthun, A. (1992). Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric Fv fragments with high avidity in *Escherichia coli*. *Biochemistry*, **31**, 1579-1584.
- Pack, P., Müller, K., Zahn, R. & Plückthun, A. (1995). Tetravalent miniantibodies with high avidity assembling in *Escherichia coli*. *J. Mol. Biol.* **246**, 28-34.
- Pelletier, J. N., Campbell-Valois, F. X. & Michnick, S. W. (1998). Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl Acad. Sci. USA*, **95**, 12141-12146.
- Pelletier, J. N., Arndt, K. M., Plückthun, A. & Michnick, S. W. (1999). An in vivo library-versus-library selection of optimized protein-protein interactions. *Nature Biotechnol.* **17**, 683-690.
- Plückthun, A. & Pack, P. (1997). New protein engineering approaches to multivalent and bispecific antibody fragments. *Immunotechnology*, **3**, 83-105.
- Potekhin, S. A., Medvedkin, V. N., Kashparov, I. A. & Venyaminov, S. (1994). Synthesis and properties of the peptide corresponding to the mutant form of the leucine zipper of the transcriptional activator GCN4 from yeast. *Protein Eng.* **7**, 1097-1101.
- Ransone, L. J. & Verma, I. M. (1990). Nuclear proto-oncogenes fos and jun. *Annu. Rev. Cell Biol.* **6**, 539-557.
- Richardson, J. S. & Richardson, D. C. (1988). Amino acid preferences for specific locations at the ends of alpha helices. *Science*, **240**, 1648-1652.
- Thompson, K. S., Vinson, C. R. & Freire, E. (1993). Thermodynamic characterization of the structural stability of the coiled-coil region of the bZIP transcription factor GCN4. *Biochemistry*, **32**, 5491-5496.
- Virnekäs, B., Ge, L., Plückthun, A., Schneider, K. C., Wellnhofer, G. & Moroney, S. E. (1994). Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucl. Acids Res.* **22**, 5600-5607.
- Weissenhorn, W., Calder, L. J., Dessen, A., Laue, T., Skehel, J. J. & Wiley, D. C. (1997). Assembly of a rod-shaped chimera of a trimeric GCN4 zipper and the HIV-1 Gp41 ectodomain expressed in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **94**, 6065-6069.
- Woolfson, D. N. & Alber, T. (1995). Predicting oligomerization states of coiled coils. *Protein Sci.* **4**, 1596-1607.
- Yu, Y., Monera, O. D., Hodges, R. S. & Privalov, P. L. (1996). Investigation of electrostatic interactions in two-stranded coiled-coils through residue shuffling. *Biophys. Chem.* **59**, 299-314.
- Zeng, X., Zhu, H., Lashuel, H. A. & Hu, J. C. (1997). Oligomerization properties of GCN4 leucine zipper e and g position mutants. *Protein Sci.* **6**, 2218-2226.
- Zhou, N. E., Kay, C. M. & Hodges, R. S. (1994). The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Eng.* **7**, 1365-1372.

Edited by A. R. Fersht

(Received 22 July 1999; received in revised form 3 November 1999; accepted 5 November 1999)