Supplementary Information

for

From DARPins to LoopDARPins: Novel LoopDARPin design allows the selection of low picomolar binders in a single round of ribosome display

Johannes Schilling, Jendrik Schöppe and Andreas Plückthun*

Biochemisches Institut
Universität Zürich
Winterthurerstrasse 190
CH-8057 Zürich, Switzerland

*Corresponding author

Correspondence should be addressed to
Prof. Andreas Plückthun, Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland, phone: +41-44-635 5570, plueckthun@bioc.uzh.ch.

Supplementary Results

Identification of a Loop consensus sequence through sequence databases

As a first step in the LoopDARPin design a loop of 10-20 amino acids in length had to be devised that maintained the integrity of the framework. Inspired by nature, we aimed at creating a loop with a stabilizing defined stem, as in the antibody CDR-H3 loop and in natural ankyrin repeat proteins that possess loop insertions [1-8]. Since sequence and structure conservation among ARs is high, we employed a consensus design strategy to identify a loop consensus sequence composed of fixed stabilizing framework positions and randomized potential interaction positions.

Structural considerations of the DARPin fold led us to the conclusion that the most promising position for the insertion of a longer loop into the existing DARPin scaffold were the protruding β-turns (cf. Figure 1 in the main text). It is at these positions that a protruding loop would have a high probability to contact the target.

To accumulate enough sequence and structural information for our loop consensus design, we first identified lead structures for further sequence analyses using the 3D structure of DARPin 2XEE [9] in a BLAST search against the PDB. Out of the 83 resulting hits, 9 natural AR structures had an insertion of varying length at one or more β-turns (ARF-GAP (PDB-ID: 1DCQ), Myotrophin (1MYO), RNaseL (1WDY), Notch (2FO1), TRPV1 (2PNN), TRPV6 (2RFA), and TRPV2 (2ETB, 2ETC [derived from the rat variant] and 2F37 [derived from the human variant]). Since the rat and human TRPV2 variants only differed by 2 amino acids outside the region of interest, the human variant was used for further analyses as a representative of this group. To obtain more distantly related structures, the 7 hit structures served as query structures in a second BLAST search against the PDB. Many hits were retrieved. However, no additional hits with an insertion at the desired position were identified.

Structural investigations of all 7 lead structures revealed that each of them could in principle serve as a separate starting point for a consensus loop design. Therefore, the sequence of each of the obtained 7 hit structures was subsequently used in a separate BLAST search against GenBank to retrieve individual sequence data sets for a consensus sequence analysis step and to delineate the natural variety of the particular ankyrin protein.

To focus the BLAST search on the β-turn region, but at the same time include a proper amount of natural AR repeat sequence in our search, we decided to use for each

lead sequence a minimal search sequence including the helix1-helix2-turn-helix1-helix2 motive (cf. Figure 3b in the main text).

With 86 sequences, the highest number of BLAST hits was retrieved for the 2F37 lead sequence. At this point, this lead structure appeared to be most promising for our consensus design approach, since it had both the most structural and sequence data available of all leads. The 86 BLAST hits obtained for the 2F37 lead sequence were then manually aligned and analyzed, yielding Loop consensus A. The 2F37 lead sequence possesses an 18 amino acid insertion at the investigated β-turn position. The performed search resulted in hits that possessed an insertion of up to 19 amino acids (as the 2PNN lead sequence, see below). Thus, the Loop consensus A sequence has a length of 19 amino acids. Sequence conservation at the beginning and the end of the insertion is very high. Diversity of the insertion regarding the occurring amino acids and, to some extent length, is confined to the middle of the loop insertion (cf. Figure 3a in the main text).

Refinement of Loop consensus A using structural data

Positions in the Loop consensus A crucial for the structural integrity of the loop and positions suitable for randomization needed to be identified. We therefore extended our sequence database analysis and included structural data in our final refinement, leading to Loop consensus B (cf. Figure 3a and Figure 3b in the main text).

The 2F37 lead structure is that of a natural AR, which is part of the conserved N-terminal intracellular domain of all transient receptor potential (TRP) channels of the vanilloid receptor subfamily (TRPV). Structures have been solved for TRPV1 (2PNN) [5], rat TRPV2 (structures from 2 crystal forms: 2ETB and 2ETC) [6], human TRPV2 (2F37) [7] and TRPV6 (2RFA) [8] (cf. Figure 3b in the main text). Since sequences of all of these structures were included in the 86 BLAST hits, yielding Loop consensus A, and since they all possess insertions of similar length at equal positions of the β-turn (Length of insertion: 2PNN: 19 amino acids and 2ETB/2ETC, 2F37 and 2RFA: 18 amino acids) structural information from all of these structures was used in our design refinement.

All 5 structures were structurally aligned with the original DARPin 2XEE using InsightII. It became immediately apparent that the aligned hit structures all share a structurally conserved stem and a structurally diverse tip of the loop. This tip varies slightly in length, with 2PNN possessing one amino acid more in this flexible region than the other structures (cf. Figure 3b in the main text). We decided to include a maximal

number of loop residues in our loop consensus design. The longest loop is present in 2PNN. Therefore, the subsequent position numbering corresponds to the 2PNN loop length.

The conserved hydrophobic stem of the loop consists of positions 1 to 3, which engage positions 6, 7 and 16 to 19 through a network of H-bonds and hydrophobic interactions. Positions 6 and 7 maintain positions 4 and 5 in a rather conserved conformation, although position 4 and 5 are not involved in stabilizing contacts and instead point towards the solvent. Together with positions 8 to 15, they are potential target interaction residues, since they appear to be either conformationally flexible or point away from the DARPin body (cf. Figure 3b in the main text).

Framework loop positions

In the structurally conserved stem of the loop, Ala1 (cf. numbering in Figure 3a in the main text) makes a H-bond to Glu19 and hydrophobic interactions to Phe6 (cf. Figure 3b in the main text). The importance of Ala at this position is reflected by the fact that it is the only amino acid present at this position in all structures and sequences. Although no clear preference for position 2 could be derived from sequence, Thr was chosen for this position, since it left Ala1, a connector between the new loop and the original DARPin framework, in a conformation closest resembling the original DARPin structure. Thr2 is further involved in a backbone H-bond with the Gln7 side chain. Furthermore, position 3 is located next to the first potential randomized loop position and thus might need to adopt many different conformations. This is therefore accommodated best by Gly, as reflected by its high level of conservation in structure as well as in sequence. Phe6 is highly conserved and makes hydrophobic interactions to Ala1. Gln7 is not conserved from the sequence analysis, but points towards the backside of the DARPin, away from a potential target and makes an H-bond to Thr2. Therefore, it was included as a framework position. Tyr16 and Phe17 are the first two of 4 conserved residues connecting the loop back to the main scaffold and are involved in the hydrophobic packing of the loop. Gly18 serves again as a flexible linking residue, which ensures the correct positioning of Glu19. Glu19 is involved in a highly conserved H-bond with Ala1 and threads the loop back into the original DARPin structure (cf. Figure 3a and Figure 3b in the main text).

Potential target interaction loop positions

Of position 4 and 5, only position 5 is conserved in sequence. Both positions, however, are not involved in stabilizing interactions and point away from the DARPin. Therefore, they were chosen to be randomized. Positions 8 to 15 are highly flexible and point towards a possible target. Together with the fact that they are not conserved on the sequence level, they were chosen as randomized positions. Although Phe15 is highly conserved, it was chosen to be randomized, since Phe at this position would shield large parts of the DARPin paratope (cf. Figure 3a and Figure 3b in the main text).

Assembly of LoopDARPin DNA libraries

Novel design features were introduced in our LoopDARPin library also on the DNA level. At the end of each repeat and surrounding the loop, unique restriction endonuclease recognition sites were introduced without changing the amino acid sequence, allowing a quick exchange of modules (Figure 1 and Suppl. Figure S1), unlike in the original design [10]. This would also permit the subsequent creation of different loop libraries, e.g. with different lengths, and would greatly facilitate the amplification of specific modules in a full-length LoopDARPin, due to repeat-specific primer annealing sites.

Assembly of capping repeats

Randomized versions of capping repeats (N-cap: N_{ran} and C-cap: C_{ran}) as well as a modified version of the previously described C-capping repeat present in the original DARPin library (C_{old}) [10,32] were PCR assembled and mixed with unrandomized capping repeats N and C. The unrandomized N-cap was mixed with the randomized library N_{ran} in a 2:1 ratio to yield a N-cap library mixture (N_{mix}). The modified version of the previously described C-capping repeat present in the original DARPin library (C_{old}), the stabilized C-cap version (C) and the randomized C-cap library (C_{ran}) were mixed in a 1:1:2 molar ratio, resulting in a C-cap library mixture (C_{mix}) (Suppl. Figure S1 and Suppl. Figure S2). Our modified version of the previously described C-capping repeat present in the original DARPin library (C_{old}) represents a hybrid between the C-capping repeat from the original DARPin library without loop [10,32] and the stabilized C-cap version "C" [32]. The concave

side of our hybrid C-capping repeat, which could potentially interact with a target, was adopted from the C-capping repeat present in the original DARPin library without loop [10,32]. For helix 2 at the convex (outer) side of our modified version of the C-capping repeat C_{old}, stabilizing residues present in the stabilized C-cap version "C" [32] were included (see Suppl. Materials and Methods and Suppl. Figure S1 and Suppl. Figure S2). In 91% (32 of 35 sequences) of the N_{mix}-repeats and in 96% (27 of 28 sequences) of the C_{mix}-repeats no frame-shift mutations were present. The N_{mix} library contained 66% N and 34% N_{ran}. The C_{mix} library contained, as intended, 25% C_{old}, 28% C and 47% C_{ran}.

Assembly of full-length ribosome-display format LoopDARPin libraries

The designed internal repeats were ligated stepwise by using type IIs restriction enzymes [10] to yield the three repeat I-IL-IF core library (Suppl. Figure S2). For ribosome-display selections, the naïve LoopDARPin library requires additional flanking sequence features present in the pRDV ribosome-display vector, such as the promoter and Shine-Dalgarno sequence upstream, and the tether and RNA hairpin downstream [11,12]. Therefore, the LoopDARPin library is ligated into and PCR-amplified out of the ribosome-display vector in each ribosome display round.

An efficient way to add either capping repeats with ensured correct sequence or different cap variants to the I-IL-IF core library is therefore to include the caps already in the pRDV vector. To make this possible, restriction endonuclease recognition sites necessary for the insertion of the core library between the caps, which will then be present on the vector, must be introduced, but to make them unique sites for these enzymes, first some sites had to be deleted from the pRDV plasmid backbone (see Supplementary Materials and Methods).

Finally we created two separate capping repeat containing pRDV vectors. pRDV1 contained unrandomized caps N and C, whereas pRDV3 contained N_{mix} and C_{mix} (cf. main text). With this strategy, the generic fusion of either unrandomized or randomized caps to our core LoopDARPin library was possible in a single step (Suppl. Figure S2). From the estimated number of molecules after each ligation step, the practical library size of both libraries was approximately 10^{11}.

Theoretical diversities of single repeats were N_{ran}-repeat $17^2 = 2.9 \cdot 10^2$, I-repeat $17^6 = 2.4 \cdot 10^7$, IL-repeat $17^3 \cdot 18^{10} = 1.8 \cdot 10^{16}$, IF-repeat $17^6 = 2.4 \cdot 10^7$ and C_{ran}-repeat $17^5 = 1.4 \cdot 10^6$. We denote the domain composition of the LoopDARPin library with 3

internal repeats as N-I-IL-IF-C (for nonrandomized caps), or LD_N3C for short, and the one with randomized caps as N_{mix}-I-IL-IF-C_{mix}, or LD_N_{mix}3C_{mix} for short. LoopDARPin libraries thus have a theoretical diversity of $1 \cdot 10^{31}$ and $4.2 \cdot 10^{39}$, respectively, vastly exceeding the practical diversity of all display systems.

Sequence analyses of unselected library members

The observed amino acid distribution of randomized internal and capping repeats was further compared to the expected distribution. In total, 1182 non-loop positions and 960 loop positions were sequenced for internal and capping repeats. Overall, observed and expected distributions were very similar. At non-loop positions, Trp was overrepresented (7.3% observed, compared to 4.3% expected), and Asn and Arg were underrepresented (4.2% and 4.1% observed, compared to 7% expected). The remaining 14 codons were found at a frequency less than 2% different from the expected value. In loop positions, Tyr and Val were overrepresented (31.4% and 4.1% observed, compared to 25% and 2.5% expected, respectively). Only Ser, Gly and Arg were underrepresented (23.6%, 4.5% and 1.5% observed, as compared to 25%, 10% and 2.5% expected, respectively). For the other 14 codons, expected and observed values differed by less than 1%. No codons other than the ones included in the randomization strategy were found. In rare cases entire trinucleotides were missing (I- and IF-repeat, 0.14%) or added (IL-repeat, 0.49%).

Evaluation of the cap randomization strategy

To determine if certain cap variants were strongly enriched at the expense of others during selections, their occurrence in the selected LoopDARPins was compared to the cap distribution present in the naïve LD_N_{mix}3C_{mix} LoopDARPin library. The naïve library was designed to contain cap variants in the following ratios: N_{mix} = N 67%, N_{ran} 33% and C_{mix} = C_{old} 25%, C 25%, C_{ran} 50%. If neither of the variants had an evolutionary advantage, this percentages should also be present in the selected pools. The distribution of cap variants in initial hits (26 binders), where the screening criterion was a specific ELISA signal over background, as well as in final hits (13 binders), where binders had to pass additional requirements such as monomeric SEC behavior and slow off-rates during SPR

screening, was evaluated. However, due to the limited number of available sequences for selected LoopDARPins it is difficult to draw statistically well-supported conclusions and only tendencies concerning enrichment of distinct cap variants can be given. Suppl. Figure 14 shows the observed occurrence of certain cap variants and the combination of N- and C-cap variants in both initial and final hit binders in comparison to the expected occurrence derived from the distribution in the naïve library.

Although N_{ran} is slightly enriched over the N variant of the N-cap in initial hits (N: 67% in the naïve library versus 54% in initial hits and N_{ran}: 33% in the naïve library versus 46% in initial hits), both the occurrence of N and N_{ran} in final hits is practically the same as in the naïve library (N: 67% in the naïve library versus 62% in the final hits and N_{ran}: 33% in the naïve library versus 38% in the final hits).

The C_{old} variant is slightly enriched at the expense of C_{ran} in initial hits, but not in final hits. The C variant of the C-cap is significantly enriched at the expense of C_{ran} in both initial and final hits. 38% of the initial and 23% of the final hits posses the C_{old} variant (25% in the naïve library). 38% of all initial and 54% of all final hits possess the C variant of the C-cap (25% in the naïve library). Only 23% of the initial and final hits contain the randomized C-cap variant C_{ran} (50% in the naïve library). Interestingly, fewer binders with the C_{old} variant met all screening criteria to become final hits, whereas almost all initial hits containing the C variant are eventually present in the final hits. Moreover, the fraction of binders with randomized C-cap (C_{ran}) in initial and final binders is the same (23%). However, in total, they are considerably less abundant than in the naïve library, where the C_{ran} variant constituted 50% of all C-cap variants.

More binders from selections against more targets are needed to really draw statistically supported conclusions concerning the advantage or disadvantage of certain cap variants and cap combinations.

The randomized C-cap C_{ran} appears less frequently in both initial and final hits, compared to its occurrence in the naïve library. Even though we showed that Cap randomization has no negative effect on the stability of LoopDARPins, the observed increased oligomerization tendency may be unfavorable in the selection process. Conversely, this finding suggests that monomeric behavior is a selected trait. In those cases where C_{ran} is selected, it remains in the final hits, and is thus beneficial in these clones. Therefore, to improve the balance between self-complementarity and the ability to provide a continuous target interaction surface, future fine-tuning of C-cap libraries might be necessary. Interestingly, the N_{ran}-C combination is significantly enriched in both initial

and final hits, illustrating the advantage of N-cap randomization, provided it is not further challenged by C_{ran} but "balanced" by C.

Most importantly, all cap variants and combinations (except for the N_{ran}-C_{old} combination, which is present in low frequency in the naïve library anyway) are present in the final hits. This indicates that having all combinations available is clearly beneficial and underlines the success of our design strategy of providing addition diversity in the capping repeats.

Characterization of LoopDARPin/target interactions by surface plasmon resonance (SPR), using cognate and non-cognate targets

SPR experiments were conducted to determine the K_D of the interactions between selected LoopDARPins and the target they have originally been selected against (cf. Figure 7c in the main text and Suppl. Figure S10c). Moreover, SPR was used to examine cross-reactivity of selected LoopDARPins with related or non-related target proteins. In total 55 LoopDARPin/target interactions were investigated (Suppl. Figure S11). The K_Ds determined by SPR for interactions between all targets and all LoopDARPins/DARPins are summarized in Figure 8 in the main text. The kinetic data of the interactions could be well evaluated with a global fit using the Langmuir [13] model. The binding of 12 out of all 14 selected binders to the target they have initially been selected against ("cognate target") can be well described by a 1:1 Langmuir binding model. Only the 4.3.E_E7/ERK2 and the 003_C9/BCL-W interaction were not fitted well by a 1:1 Langmuir binding model (see Figure 8 in the main text).

Our selection pressure focused on high affinity and groove binding, but not on specificity for a particular family member over all other family members. Thus, selections against BCL-2, BCL-XL and BCL-W yielded some binders, which did not only bind to the target they had been selected against (the "cognate" target), but also to one or two of the three remaining — closely related — family members (none of the BCL-2, BCL-XL and BCL-W binders bound to MCL-1 and vice versa), albeit with distinctly lower affinity. Interestingly, if binding to other family members was detected, in approximately 50% of these cases, this binding was not fitted well by a 1:1 Langmuir binding model (Figure 8 in the main text and Suppl. Figure S11).

Due to the described conformational flexibility of the loop in LoopDARPins, the heterogeneous analyte model (the analyte being the LoopDARPin in the set-up used) was

tested to fit the data, but could not describe the observed kinetic data of the interaction. However, the binding was fitted well by the heterogeneous ligand model. Nonetheless, the goodness-of-fit alone cannot be used to justify the use of a certain model [14,15]. The justification of a used model must come from other experimental data, which prompted us to exclude the heterogeneous ligand model, since it conflicted with some observations, which will now be described in more detail.

The heterogeneous ligand model assumes that one analyte binds two different ligand species. Because each ligand-receptor complex should have distinct affinities, the resulting sensorgram reflects the sum of two independent reactions with two sets of rate constants. However, we hypothesize that the observed two-step binding does not originate from the LoopDARPin binding to two different ligand target species. We mainly observe this two-step binding within the structurally highly homologous BCL-2, BCL-XL and BCL-W targets (cf. Figure 6 in the main text and Suppl. Figure S7). We therefore presume that LoopDARPins, which bind to more than one of these family members, bind to similar epitopes on the respective targets, which are, however, not identical.

Subtle structural differences of the epitope among related targets within the family may lead to the fact that LoopDARPin and non-cognate target fit less well, compared to the cognate target. Nevertheless, initial binding, mediated by loop or non-loop LoopDARPin positions, may still occur. The conformationally flexible LoopDARPin (especially in the loop region) may subsequently adapt to the target within the complex. This leads to a conformational change ("induced fit") with improved binding.

This hypothesis is supported by the fact that the interaction of LoopDARPins with the cognate target (diagonal interactions, framed in magenta in the K_D matrix of Figure 8 in the main text) can in over 90% of all cases (12 out of 13 interactions) be well described by a 1:1 Langmuir binding model (Suppl. Figure S11), and this interaction is characterized by the highest affinity. Six of 13 (46%) selected BCL-2 family binding LoopDARPins are highly specific for one target. The remaining 7 selected binders (54%) also bind to other family members.

In approximately 50% of the cases, where interaction with other family members occurs, we observe non-Langmuir behavior. In all cases, binding to other non-cognate family members is weaker than the binding to the cognate target (Suppl. Table ST2).

In rare cases, we also observed non-Langmuir behavior for LoopDARPin interactions to the target the LoopDARPin was originally selected against (2 out of 14 LoopDARPins: 4.3.E_E7/ERK2 and 003_C9/BCL-W interaction). We assume that this two-step binding is also caused by the just described mechanism of an induced fit.

Accounting for the described observations, we chose a two-step (isomerization) model [14] for the analysis of those interactions, which could not be well described by a 1:1 Langmuir binding model. The two-step (isomerization) model is the most simplified scheme by which one can quantitatively describe a two-step binding process:

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} (AB)^* \quad \text{[eq. 1]}$$

where A and B represent LoopDARPin and target and AB and (AB)* represent initial complex and final docked (rearranged) complex, respectively. (The parenthesis denotes that the whole complex has undergone a conformational change denoted by the asterisk). This model is an induced fit model, where a conformational change occurs *after* complex formation [16]. The equilibrium constants of the individual steps are

$$K_{D1} = \frac{k_{off}}{k_{on}} \quad \text{and} \quad K_2 = \frac{k_{-2}}{k_2} \quad \text{[eq. 2]}$$

and the overall equilibrium binding constant is calculated as [17,18]

$$K_D = \frac{1}{\frac{1}{K_{D1}} \cdot (1 + \frac{1}{K_2})} \quad \text{[eq. 3]}$$

Together with Figure 8 in the main text, Suppl. Table ST2 shows the K_Ds determined by SPR for interactions between all targets and all LoopDARPins/DARPins. The association rates of the interactions fitted by 1:1 Langmuir binding model and by the two-step model (k_{on}) are in the typical range for protein-protein interactions (10^4 to 10^6 $M^{-1}s^{-1}$) [19] and the dissociation rates (k_{off}) range vary from 10^{-4} to 10^{-6} s^{-1} (10^{-2} s^{-1} when including also the measurements on non-cognate targets). For LoopDARPins, which interact with more than one BCL-2 family member, the interaction with the original target is always the strongest.

For those cases where binding follows the two-step model, the initial interaction, described by K_{D1}, is always one to three orders of magnitude weaker than the interaction between LoopDARPin and its original target, which is described by a single K_D (Suppl. Table ST2) (as it is well described by a Langmuir interaction). The conversion of the

initial complex AB to (AB)* is always very slow, with rate constants k_2 between 10^{-3} to 10^{-4} s^{-1}, most likely due to a high activation energy for the conformational change to reach the transition state. However, once (AB)* is formed, it is very stable (k_{-2}: 10^{-3} to 10^{-5} s^{-1}). The equilibrium between AB and (AB)* is thus on the side of (AB)*. This means that if the initial complex is given sufficient time to stabilize through a conformational change, the overall LoopDARPin/target interaction becomes stronger. As ELISA experiments reveal (Figure 7a in the main text and Suppl. Figure 10a), LoopDARPins showing a two-step binding to related family members can still be used to discriminate between members of the family, provided concentrations of LoopDARPin below K_{D1} are used. These LoopDARPins may thus be used to specifically detect single family members in e.g. ELISA experiments.

The described findings will be briefly illustrated using the two selected LoopDARPins 008_C6 and 003_D9 as an example. They show the symmetrical nature of the obtained interaction results, where interactions of a LoopDARPin with its original target can be described by a 1:1 Langmuir binding model, whereas interactions of the same LoopDARPin with structurally highly related family members can be well explained by a two-step binding model.

Sensorgrams which show the interaction of LoopDARPins with the cognate target are highlighted by a pink frame in Suppl. Figure S11. LoopDARPin 008_C6 originates from the one-round selection against BCL-2, and its interaction with BCL-2 can be well described by a 1:1 Langmuir binding model. The interaction possesses a K_D of 30 pM with rate constants $k_{on} = 1.56 \cdot 10^5$ $M^{-1}s^{-1}$ and $k_{off} = 4.64 \cdot 10^{-6}$ s^{-1}. Interactions to the structurally highly related BCL-XL and BCL-W show a two-step binding and can be well described by a two-step (isomerization) model. For the 008_C6/BCL-XL interaction, the $K_{D1} = 3.10 \cdot 10^{-8}$ M ($k_{on} = 8.98 \cdot 10^4$ $M^{-1}s^{-1}$ and $k_{off} = 2.79 \cdot 10^{-3}$ s^{-1}) is three orders of magnitudes greater (worse) than for the 008_C6/BCL-2 interaction. Together with $K_2 = 3.38 \cdot 10^{-2}$ ($k_2 = 5.38 \cdot 10^{-4}$ s^{-1} and $k_{-2} = 1.82 \cdot 10^{-5}$ s^{-1}), the 008_C6/BCL-XL interaction has an overall K_D of $1.01 \cdot 10^{-9}$ M, as calculated according to Eq. [3].

For the 008_C6/BCL-W interaction, the $K_{D1} = 3.12 \cdot 10^{-8}$ M ($k_{on} = 1.07 \cdot 10^5$ M^{-1} s^{-1} and $k_{off} = 3.34 \cdot 10^{-3}$ s^{-1}) is also three orders of magnitudes greater (worse) than for the 008_C6/BCL-2 interaction and, together with $K_2 = 1.03 \cdot 10^{-1}$ ($k_2 = 7.47 \cdot 10^{-4}$ s^{-1} and $k_{-2} = 7.71 \cdot 10^{-5}$ s^{-1}), the 008_C6/BCL-W interaction has an overall K_D of $2.92 \cdot 10^{-9}$ M.

LoopDARPin 003_D9 originates from the three-round selection against BCL-W. Its interaction to BCL-W, the cognate target, can be well described by a 1:1 Langmuir binding model. For this interaction, a K_D of 63 pM can be derived with rate constants $k_{on} =$

$7.42 \cdot 10^4$ $M^{-1}s^{-1}$ and $k_{off} = 4.70 \cdot 10^{-6}$ s^{-1}. Interactions to the structurally highly related BCL-2 and BCL-XL show a two-step binding and can be well described by a two-step model. For the 003_D9/BCL-2 interaction the $K_{D1} = 2.30 \cdot 10^{-8}$ M ($k_{on} = 4.79 \cdot 10^4$ $M^{-1}s^{-1}$ and $k_{off} = 1.10 \cdot 10^{-3}$ s^{-1}) is three orders of magnitudes greater (worse) than for the 003_D9/BCL-W interaction. The 003_D9/BCL-2 interaction has an overall K_D of $7.85 \cdot 10^{-9}$ M with $K_2 = 5.19 \cdot 10^{-1}$ ($k_2 = 1.62 \cdot 10^{-3}$ s^{-1} and $k_{-2} = 8.40 \cdot 10^{-4}$ s^{-1}). For the 003_D9/BCL-XL interaction, the $K_{D1} = 8.15 \cdot 10^{-8}$ M ($k_{on} = 9.68 \cdot 10^4$ $M^{-1}s^{-1}$ and $k_{off} = 7.89 \cdot 10^{-3}$ s^{-1}) is also three orders of magnitudes greater (worse) than for the 003_D9/BCL-W interaction and, together with $K_2 = 7.16 \cdot 10^{-2}$ ($k_2 = 2.96 \cdot 10^{-3}$ s^{-1} and $k_{-2} = 2.12 \cdot 10^{-4}$ s^{-1}), the 003_D9/BCL-XL interaction has an overall K_D of $5.45 \cdot 10^{-9}$ M.

Supplementary Materials and Methods

Generation of combinatorial LoopDARPin libraries

Standard oligonucleotides were from Microsynth AG (Switzerland) and oligonucleotides possessing mixed trinucleotides as building blocks [20] were from Metabion International AG (Germany). Single N, I, IL, IF, C_{old} and C repeats were generated using assembly PCR with Vent® Polymerase (1 min annealing at 50°C and standard buffers). Vent exo-® Polymerase was used for N_{ran}- and C_{ran}-repeat assembly. Oligonucleotides used for assembly PCR of single repeats are listed in Suppl. Table ST3. The LoopDARPin modules were generated using the following oligonucleotides: For the N-cap: TEN3, EWT1, EWT2N, INT6. For the internal repeat I: INT1, INT2, INT3I, INT4, INT5I_N, INT6aI. For the loop-containing internal repeat IL: INT1L, INT2L, INT3L, INT4L INT5, INT6aL. For the loop-following internal repeat IF: INT1F, INT2, INT3F, INT4, INT5, INT6aF_C. For the modified version of the previously described C-capping repeat present in the original DARPin library C_{old}: INT5, WTC1, WTC2C, WTC3Cno. For the stabilized C-cap version C: INT5, WTC1C, WTC2C, WTC3Cno. For the randomized N-cap N_{ran}: EWT1Nran, EWT1Nran0, EWT1Nran00, EWT1Nran000, EWT2N, EWT2N0, EWT2N00, EWT2N000, INT2, INT6. For the randomized C-cap C_{ran}: INT1, INT2, INT4CCapranfwd, INT4CCapranrev, INT5Bpu10I, WTC2Cran2, WTC3CnoWTC2Cran2. For sequence analyses a subset of the resulting PCR products was cloned via *Bam*HI/*Hin*dIII into pPANK. LoopDARPin module sequences are shown in Suppl. Figure S1. The ligation of single modules to yield the I-IL-IF core library was performed as described in ref. [10].

Generation of novel ribosome-display vectors (pRDV1 and pRDV3)

pRDV1 and pRDV3 were derived from the ribosome display vector pRDV [11]. Starting from pRDV, the *Bst*API restriction sites (position 1160 and 1428 in the nucleotide sequence) and the *Apo*I restriction site at position 968 were mutated using the QuikChange® Multi Site-Directed Mutagenesis Kit (Stratagene, La Jolla, USA) with oligonucleotides pRDVqc1, pRDVqc2 and pRDVqc3. *Apo*I restriction sites at position 2425 and 2436 were removed through *Apo*I digest, mung bean nuclease digestion and blunt end ligation. To introduce unrandomized caps, a LoopDARPin in pPANK was

PCR-amplified using oligonucleotides EWT5k and WTC3Cno. The resulting PCR product was *Nco*I/*Hin*dIII digested and ligated to the *Nco*I/*Hin*dIII digested mutated pRDV vector. The resulting vector pRDV0 served as a starting point to generate pRDV1 and pRDV3. To generate pRDV1, oligonucleotides bLA_BstAPI and bLA_ApoI were used in a PCR using pPANK as a template. The resulting PCR product was *Bst*API/*Apo*I digested and ligated to the *Bst*API/*Apo*I digested pRDV0 vector to yield pRDV1. pRDV3 was generated using oligonucleotides bLA_BstAPI and bLA_ApoI_Bpu10I in a PCR using pPANK as a template. The resulting PCR product was *Bst*API/*Bpu*10I digested and ligated to the *Bst*API/*Bpu*10I digested pRDV0 to yield pRDV2. Stepwise, the N-cap library mixture N_{mix} and the C-cap library mixture C_{mix} were introduced into pRDV2 through *Bam*HI/*Bst*API and *Bpu*10I/*Hin*dIII digest and ligation, respectively, to generate pRDV3.

Generation of antigen production vectors (pAT222-pD, pAT223-pD)

pAT222-pD and pAT223-pD were derived from pAT222 and pAT223 [21,22], respectively. To generate pAT222-pD, oligonucleotides pAT222to-pD_f and pAT222to-pD_r were used in a PCR reaction. The resulting PCR product was cloned via *Eco*RI/*Bam*HI into pAT222. pAT223-pD was generated through annealing of oligonucleotide pAT223to-pD_f and pAT223to-pD_r and ligation of the annealed oligonucleotides into *Nco*I/*Bam*HI digested pAT223. The sequence of the oligonucleotides used can be found Suppl. Table ST5.

References

1. Morea, V., Tramontano, A., Rustici, M., Chothia, C. & Lesk, A.M. (1998). Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J. Mol. Biol.* **275**, 269–294
2. Shirai, H., Kidera, A. & Nakamura, H. (1996). Structural classification of CDR-H3 in antibodies. *FEBS Lett.* **399**, 1–8
3. Kuroda, D., Shirai, H., Kobori, M. & Nakamura, H. (2008). Structural classification of CDR-H3 revisited: a lesson in antibody modeling. *Proteins* **73**, 608–620
4. Choi, Y. & Deane, C.M. (2011). Predicting antibody complementarity determining region structures without classification. *Mol. Biosyst.* **7**, 3327–3334
5. Lishko, P.V., Procko, E., Jin, X., Phelps, C.B. & Gaudet, R. (2007). The ankyrin repeats of TRPV1 bind multiple ligands and modulate channel sensitivity. *Neuron* **54**, 905–918
6. Jin, X., Touhey, J. & Gaudet, R. (2006). Structure of the N-terminal ankyrin repeat domain of the TRPV2 ion channel. *J. Biol. Chem.* **281**, 25006–25010
7. McCleverty, C.J., Koesema, E., Patapoutian, A., Lesley, S.A. & Kreusch, A. (2006). Crystal structure of the human TRPV2 channel ankyrin repeat domain. *Protein Sci.* **15**, 2201–2206
8. Phelps, C.B., Huang, R.J., Lishko, P.V., Wang, R.R. & Gaudet, R. (2008). Structural analyses of the ankyrin repeat domain of TRPV6 and related TRPV ion channels. *Biochemistry* **47**, 2476–2484
9. Kramer, M.A., Wetzel, S.K., Plückthun, A., Mittl, P.R.E. & Grütter, M.G. (2010). Structural determinants for improved stability of designed ankyrin repeat proteins with a redesigned C-capping module. *J. Mol. Biol.* **404**, 381–391
10. Binz, H.K., Stumpp, M.T., Forrer, P., Amstutz, P. & Plückthun, A. (2003). Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *J. Mol. Biol.* **332**, 489–503
11. Zahnd, C., Amstutz, P. & Plückthun, A. (2007). Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods* **4**, 269–279
12. Plückthun, A. (2012). Ribosome display: a perspective. *Methods Mol. Biol.* **805**, 3–28
13. Brookst, I. (1993). Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods. *Anal. Biochem. 212, 457-468*
14. Morton, T.A., Myszka, D.G. & Chaiken, I.M. (1995). Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration. *Anal. Biochem.* **227**, 176–185
15. Schuck, P. (1997). Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 541–566
16. Foote, J. & Milstein, C. (1994). Conformational isomerism and the diversity of antibodies. *Proc. Natl. Acad. Sci. USA* **91**, 10370–10374
17. Krauss, G., Riesner, D. & Maass, G. (1976). Mechanism of discrimination between cognate and non-cognate tRNAs by phenylalanyl-tRNA synthetase from yeast. *Eur. J. Biochem.* **68**, 81–93
18. Riesner, D., Pingoud, A., Boehme, D., Peters, F. & Maass, G. (1976). Distinct steps

in the specific binding of tRNA to aminoacyl-tRNA synthetase. Temperature-jump studies on the serine-specific system from yeast and the tyrosine-specific system from Escherichia coli. *Eur. J. Biochem.* **68**, 71–80
19. Wodak, S.J. & Janin, J. (2002). Structural basis of macromolecular recognition. *Adv. Protein Chem.* **61**, 9–73
20. 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. *Nucleic Acids Res.* **22**, 5600–5607
21. Binz, H.K., Amstutz, P., Kohl, A., Stumpp, M.T., Briand, C., Forrer, P., Grütter, M.G. & Plückthun, A. (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* **22**, 575–582
22. Forrer, P. & Jaussi, R. (1998). High-level expression of soluble heterologous proteins in the cytoplasm of Escherichia coli by fusion to the bacteriophage lambda head protein D. *Gene* **224**, 45–52
23. Petros, A.M., Medek, A., Nettesheim, D.G., Kim, D.H., Yoon, H.S., Swift, K., Matayoshi, E.D., Oltersdorf, T. & Fesik, S.W. (2001). Solution structure of the antiapoptotic protein bcl-2. *Proc. Natl. Acad. Sci. USA* **98**, 3012–3017
24. Feng, W., Huang, S., Wu, H. & Zhang, M. (2007). Molecular basis of Bcl-xL's target recognition versatility revealed by the structure of Bcl-xL in complex with the BH3 domain of Beclin-1. *J. Mol. Biol.* **372**, 223–235
25. Kummer, L., Parizek, P., Rube, P., Millgramm, B., Prinz, A., Mittl, P.R.E., Kaufholz, M., Zimmermann, B., Herberg, F.W. & Plückthun, A. (2012). Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. *Proc. Natl. Acad. Sci. USA* **109**, E2248–E2257
26. Denisov, A.Y., Madiraju, M.S.R., Chen, G., Khadir, A., Beauparlant, P., Attardo, G., Shore, G.C. & Gehring, K. (2003). Solution structure of human BCL-w: modulation of ligand binding by the C-terminal helix. *J. Biol. Chem.* **278**, 21124–21128
27. Lawrence, M.C., Jivan, A., Shao, C., Duan, L., Goad, D., Zaganjor, E., Osborne, J., McGlynn, K., Stippec, S., Earnest, S., Chen, W. & Cobb, M.H. (2008). The roles of MAPKs in disease. *Cell Res.* **18**, 436–442
28. Fire, E., Gullá, S.V., Grant, R.A. & Keating, A.E. (2010). Mcl-1-Bim complexes accommodate surprising point mutations via minor structural changes. *Protein Sci.* **19**, 507–519
29. Chen, Z., Gibson, T.B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C. & Cobb, M.H. (2001). MAP kinases. *Chem. Rev.* **101**, 2449–2476
30. Denisov, A. Y., Chen, G., Sprules, T., Moldoveanu, T., Beauparlant, P. & Gehring, K. (2006). Structural model of the BCL-w-BID peptide complex and its interactions with phospholipid micelles. *Biochemistry* **45**, 2250–2256
31. Green, D.R. & Evan, G.I. (2002). A matter of life and death. *Cancer Cell* **1**, 19–30
32. Interlandi, G., Wetzel, S.K., Settanni, G., Plückthun, A. & Caflisch, A. (2008). Characterization and further stabilization of designed ankyrin repeat proteins by combining molecular dynamics simulations and experiments. *J. Mol. Biol.* **375**, 837–854

Supplementary Tables

Suppl. Table ST1. Statistics for data collection and refinement

LoopDARPin	**Statistics**
$N_{ran}1$_G06_C	
Data collection	
Space group	oP: $P2_12_12_1$
Cell dimensions, Å	a = 49.6, b = 63.9, c = 113.6 α = β = γ = 90°
AU content	2 molecules
VM, $Å^3$/Da	2.27
Resolution limits, Å	50 – 1.7
Observed reflections	Total 132319, unique 40279, possible 40530
Completeness, %	99.4 (99.5)*
R-merge, %	46.7 (5.4)*
I/σ	14.15 (2.51)*
Refinement	
Resolution range, Å	45.5-1.7
Final R-cryst, R-free, %	16.6, 19.2
Number of residues	341
Number of solvent molecules	488
Number of atoms	3118
Mean B-factor, $Å^2$	21.6
rmsd (bonds), Å	0.006
rmsd (angles), °	0.991
Ramachandran analysis, %	98.3/1.7/0

* Values in parentheses refer to the highest-resolution shell
oP (orthorhombic primitive)

Suppl. Table ST2. Kinetic binding data for all selected LoopDARPins/DARPins determined by surface plasmon resonance. Where 2 sets of binding constants are shown, fits to the model of Eq 1 were used, see Suppl. Results.

LoopDARPin	Target	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_{D1} (M)	k_2 (s^{-1})	k_{-2} (s^{-1})	K_2	K_D (M)
008_H10	**BCL-2**	2.03×10^{5}	1.17×10^{-5}					5.79×10^{-11}
	BCL-XL	-	-					no binding
	BCL-W	-	-					no binding
	MCL-1	-	-					no binding
008_C6	**BCL-2**	1.56×10^{5}	4.64×10^{-6}					2.97×10^{-11}
	BCL-XL	8.98×10^{4}	2.79×10^{-3}	3.10×10^{-08}	5.38×10^{-4}	1.82×10^{-5}	3.38×10^{-2}	1.01×10^{-09}
	BCL-W	1.07×10^{5}	3.34×10^{-3}	3.12×10^{-08}	7.47×10^{-4}	7.71×10^{-5}	1.03×10^{-1}	2.92×10^{-09}
	MCL-1	-	-					no binding
008_F2	**BCL-2**	5.90×10^{4}	3.09×10^{-5}					5.23×10^{-10}
	BCL-XL	7.88×10^{4}	6.74×10^{-3}					8.56×10^{-08}
	BCL-W	-	-					no binding
	MCL-1	-	-					no binding
008_D3	**BCL-2**	2.82×10^{5}	2.22×10^{-4}					7.87×10^{-10}
	BCL-XL	3.71×10^{5}	3.81×10^{-3}	1.03×10^{-08}	7.92×10^{-4}	2.49×10^{-4}	3.14×10^{-1}	2.46×10^{-09}
	BCL-W	2.92×10^{5}	2.27×10^{-3}	7.78×10^{-09}	5.41×10^{-4}	2.97×10^{-4}	5.49×10^{-1}	2.76×10^{-09}
	MCL-1	-	-					no binding
001_C1	**BCL-2**	9.85×10^{4}	4.75×10^{-5}					4.82×10^{-10}
	BCL-XL	2.35×10^{5}	9.37×10^{-3}	3.98×10^{-08}	1.34×10^{-3}	6.51×10^{-4}	4.86×10^{-1}	1.30×10^{-08}
	BCL-W	-	-					no binding
	MCL-1	-	-					no binding
001_C10	**BCL-2**	3.40×10^{5}	6.60×10^{-5}					1.94×10^{-10}
	BCL-XL	-	-					no binding
	BCL-W	1.40×10^{6}	3.00×10^{-2}					2.31×10^{-08}
	MCL-1	-	-					no binding
011_F12	**BCL-2**	9.38×10^{4}	5.23×10^{-5}					5.57×10^{-10}
	BCL-XL	2.02×10^{5}	2.14×10^{-3}					1.06×10^{-08}
	BCL-W	2.65×10^{5}	1.86×10^{-3}					7.02×10^{-09}
	MCL-1	-	-					no binding
012_F12	**BCL-2**	7.04×10^{4}	1.02×10^{-4}					1.44×10^{-09}
	BCL-XL	1.90×10^{5}	1.97×10^{-4}					1.04×10^{-09}
	BCL-W	1.70×10^{5}	4.23×10^{-4}					2.49×10^{-09}
	MCL-1	-	-					no binding
003_D9	**BCL-2**	4.79×10^{4}	1.10×10^{-3}	2.30×10^{-08}	1.62×10^{-3}	8.40×10^{-4}	5.19×10^{-1}	7.85×10^{-09}
	BCL-XL	9.68×10^{4}	7.89×10^{-3}	8.15×10^{-08}	2.96×10^{-3}	2.12×10^{-4}	7.16×10^{-2}	5.45×10^{-09}
	BCL-W	7.42×10^{4}	4.70×10^{-6}					6.33×10^{-11}
	MCL-1	-	-					no binding
003_C9	**BCL-2**	1.85×10^{5}	2.61×10^{-3}	1.41×10^{-08}	1.15×10^{-3}	1.04×10^{-3}	9.04×10^{-1}	6.70×10^{-09}
	BCL-XL	4.23×10^{5}	1.04×10^{-3}	2.47×10^{-09}	2.24×10^{-4}	2.50×10^{-4}	1.12×10^{0}	1.30×10^{-09}
	BCL-W	3.52×10^{5}	8.86×10^{-4}	2.52×10^{-09}	4.14×10^{-4}	3.30×10^{-4}	7.97×10^{-1}	1.12×10^{-09}
	MCL-1	-	-					no binding
013_H8	**BCL-2**	-	-					no binding
	BCL-XL	1.22×10^{6}	8.32×10^{-3}					6.83×10^{-09}
	BCL-W	3.96×10^{5}	2.55×10^{-4}					6.43×10^{-10}
	MCL-1	-	-					no binding
013_D12	**BCL-2**	-	-					no binding
	BCL-XL	2.42×10^{6}	6.00×10^{-2}					2.66×10^{-08}
	BCL-W	7.07×10^{5}	7.27×10^{-3}					1.03×10^{-08}
	MCL-1	-	-					no binding
014_G9	**BCL-2**	-	-					no binding
	BCL-XL	-	-					no binding
	BCL-W	-	-					no binding
	MCL-1	1.59×10^{5}	1.23×10^{-4}					7.74×10^{-10}
4.3.E_E7	**BCL-2**	-	-					no binding
	MCL-1	-	-					no binding
	ERK2	1.75×10^{5}	1.04×10^{-3}	5.94×10^{-09}	2.39×10^{-4}	9.99×10^{-4}	4.18×10^{0}	4.79×10^{-09}

Suppl. Table ST3. Overview of oligonucleotides.

Oligonucleotide	Oligonucleotide sequence (5' to 3')
INT1	CTGACGTTAACGCTNNNGACNNNNNNGGTNNNACTCCGCTGCACCTGGC[a]
INT2	ACTCCGCTGCACCTGGCTGCTNNNNNNGGTCACCTGGAAATCG[a]
INT1L	CTGACGTTAACGCTNNNGCTACTGGTZZZZZZTTCCAGZZZZZZZZZZZZZZZZZZZZZZZZZZTACTTCGGTGAAACTCCGC TGCACCTGGC[b]
INT2L	ACTCCGCTGCACCTGGCCGCTNNNNNNGGTCACCTGGAAATCG[a]
INT1F	CTGACGTTAACGCTNNNGCTNNNNNNGGTNNNACTCCGCTGCACCTGGC[a]
TEN3	TTCCGCGGATCCG
EWT1	TTCCGCGGATCCGACCTGGGTAAGAAACTGCTGGAAGCTGCTCGTGCTGGTCAGGACGACGAAG
EWT2N	AACGTCAGCACCGTTTGCCATCAGGATACGAACTTCGTCGTCCTGACC
INT6	TTTGGGAAGCTTCTAGAAGACAACGTCAGCACCGT
INT3I	AACGTCAGCACCGGTTTTCAGCAGAACTTCAACGATTTCCAGGTGACC
INT4	AGCAGCCAGGTGCAGCGGAGT
INT5_I_N	TTCCGCGGATCCTAGCAAACGGTGCTGACGTTAACGCT
INT6aI	TTTGGGAAGCTTCTAAGGTCTCACGTCAGCACCGG
INT3L	AACGTCAGCGCCGGCCTTCAGCAGAACTTCAACGATTTCCAGGTGACC
INT4L	AGCGGCCAGGTGCAGCGGAGT
INT5	TTCCGCGGATCCTAGGAAGACCTGACGTTAACGCT
INT6aL	TTTGGGAAGCTTCTAAGGTCTCACGTCAGCGCCG
INT3F	AACGTCGGCGCCGTGCTTCAGCAGAACTTCAACGATTTCCAGGTGACC
INT6aF_C	TTTGGGAAGCTTACCGAATTTGTCCTGAGCGTTAACGTCGGCGCC
WTC1	CTGACGTTAACGCTCAGGACAAATTCGGTAAGACCGCTTTCGACATCTCCATCGACAACGGTAACGAGG
WTC1C	CTGACGTTAACGCTCAGGACAAATTCGGTAAGACTCCGTTCGACCTGGCTATCGACAACGGTAACGAGG
WTC2C	AGCAGCTTTCTGCAGAACTTCAGCGATGTCCTCGTTACCGTTGTC
WTC3Cno	TTTGGGAAGCTTAGCAGCTTTCTGCAGAACTTCAGC
EWT1Nran	TTCCGCGGATCCGACCTGGGTAAGAAACTGCTGGAAGCTGCT
EWT1Nran0	TTCCGCGGATCCGACCTGGGTAAGAAACTGCACCTGGCTGCT
EWT1Nran00	TTCCGCGGATCCGACCTGGGTAAGAAACTGCTGCTGGCTGCT
EWT1Nran000	TTCCGCGGATCCGACCTGGGTAAGAAACTGCTGGAGGCTGCT
EWT2N	AACGTCAGCACCGTTTGCCATCAGGATACGAACTTCGTCGTCCTGACC
EWT2N0	AACGTCAGCACCGTTTGCCATCAGGATACGAACTTCGTCCAGGTGACC
EWT2N00	AACGTCAGCACCGTTTGCCATCAGGATACGAACTTCGTCCAGCTGACC
EWT2N000	AACGTCAGCACCGTTTGCCATCAGGATACGAACTTCGTCGTGCTGACC
INT4CCapranfwd	ACTCCGTTCGACCTGGCTGCT
INT4CCapranrev	AGCAGCCAGGTCGAACGGAGT
INT5Bpu10I	TTCCGCGGATCCTAGGAAGACCTGACGTTAACGCTCAGGAC
WTC2Cran2	AGCAGCTTTCTGCAGAACTTCAGCGATGTCCTCGTTACC
WTC3CnoWTC2Cran2	TTTGGGAAGCTTAGCAGCTTTCTGCAGAACTTCAGCGATGTCCTCGTTACC
JSCRDif4	AGAGGATCGCATCACCATCACCATCACGGATCCGACCTGGG
WTC3Cn	TTTGGGAAGCTTCTACTAAGCAGCTTTCTGCA
T7B	ATACGAAATTAATACGACTCACTATAGGGAGACCACAACGG
TolAk	CCGCACACCAGTAAGGTGTGCGGTTTCAGTTGCCGCTTTCTTTCT
EWT5k	TTCCTCCATGGGTATGAGAGGATCG
pRDVqc1	AAAAAGCCGAGGCGGCCGCTGTGGCACTGAAGAAG
pRDVqc2	AAAGCGGCAGCGGCGAAAGCTGTAGCTGAAAAAGCCG
pRDVqc3	GGCCTCGGGGGCCGAGTACGGATCTGGTG
bLA_BstAPI	TTCCGATCCTGATGGCAAACGGTGCTGACTAAATGAGTATTCAACATTTCCGTGTCGCCC
blA_ApoI	TTTGGGAGTCTTACCGAATTTGTCGGGGTCTGACGCTCAGTGGAACG
bLA_ApoI_Bpu10I	TTTGGGAGTCTTACCGAATTTGTCCTGAGCGTTGGGGTCTGACGCTCAGTGGAACG

[a] NNN in the oligonucleotide sequence represents a mixture of trinucleotides encoding the amino acids A, D, E, H, K, N, Q, R, S, T (7% each) and F, I, L, M, V, W, Y (4.3% each).

[b] ZZZ represents a mixture of trinucleotides encoding the amino acids Y, S (25% each), G (10%), A, D, E, H, K, N, Q, R, T, F, I, L, M, V, W, P (2.5% each).

Suppl. Table ST4. Overview of oligonucleotides used for gene assembly of targets.

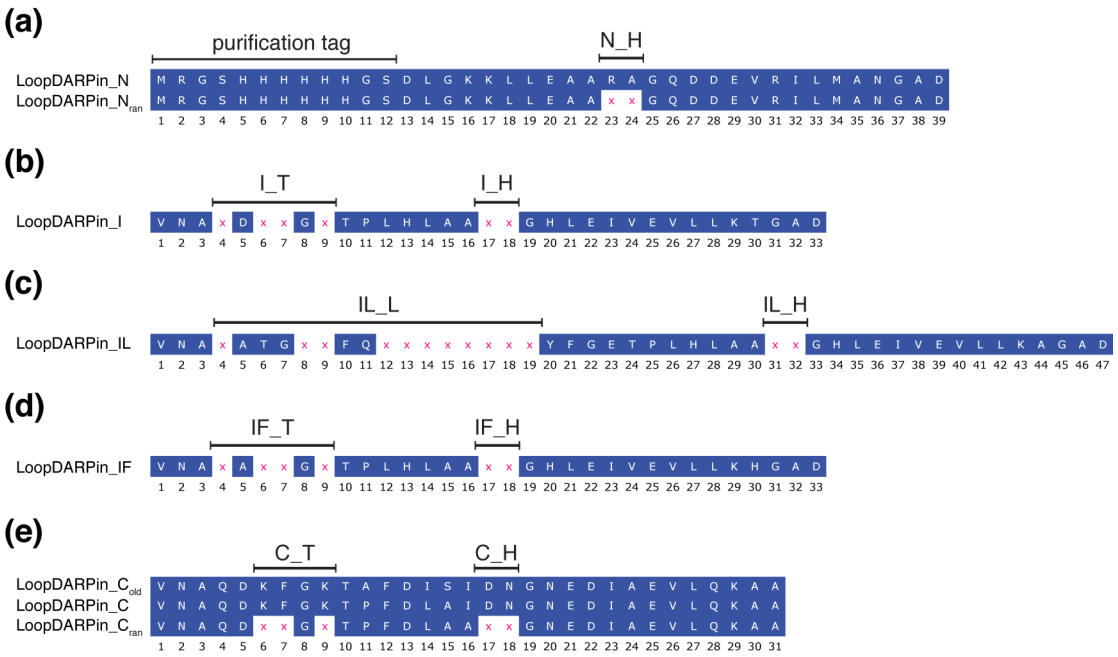
Target	Oligonucleotide	Oligonucleotide sequence (5' to 3')
ERK2	pAT223_ERK2_f	TTCCGCGGATCCATGGCGGCGGCGGCGGGCCCG
	pAT223_ERK2_r	TTTGGGAAGCTTCTACTAAGATCTGTATCCTGGCTGG
BCL-2	1G5Mbcl2_f1	GTCGTACCGGTTATGATAATCGTGAAATTGTGATGAAATATATTCATTATAAACTGAGCC
	1G5Mbcl2_f2	GATGTGGAAGAAAATCGTACCGAAGCGCCGGAAGGTACCGAAAGCGAAGTGGTGCATCTG
	1G5Mbcl2_f3	CGTCGTTATCGTGGTGATTTTGCGGAAATGAGCAGCCAGCTGCATCTGACCCCGTTTACC
	1G5Mbcl2_f4	GAACTGTTTCGTGATGGTGTGAATTGGGGTCGTATTGTGGCGTTTTTTGAATTTGGTGG
	1G5Mbcl2_f5	GTGAAATGAGCCCGCTGGTGGATAATATTGCGCTGTGGATGACCGAATATCTGAATCGTC
	1G5Mbcl2_fi1	CGTCGTTATCGTGG
	1G5Mbcl2_r1	CGATTTTCTTCCACATCATCACCCGCATCCCATTCATAACCACGCTGGCTCAGTTTATAATG
	1G5Mbcl2_r2	CCACGATAACGACGGCTAAAATCATCACCCGCCTGACGCAGCGCCAGATGCACCACTTCG
	1G5Mbcl2_r3	CCATCACGAAACAGTTCTTCCACCACGGTCGCAAAACGACCACGCGCGGTAAACGGGGTCAG
	1G5Mbcl2_r4	GCGGGCTCATTTCACGATTCACGCTTTCCACGCACATCACACCACCAAATTCAAAAAACG
	1G5Mbcl2_r5	CAAACGCATCCCAACCACCATTATCCTGAATCCAGGTATGCAGATGACGATTCAGATATTCG
	1G5Mbcl2_ri1	CCACGATAACGACG
	1G5Mbcl2_ri2	GCGGGCTCATTTCAC
	pAT223_1G5M_f	TTCCGCGGATCCCATGCGGGTCGTACCGGTTATG
	pAT222_1G5M_r	TTTGGGAAGCTTCACGCATGCTCGGACCATACAG
	pAT223_1G5M_r[a]	TTTGGGAAGCTTCTACTAACGCATGCTCGGACCATACAG
BCL-XL	2PONbclxl_f1	CGTGAACTGGTGGTGGATTTTCTGAGCTATAAACTGAGCCAGAAAGGTTATAGCTGGAGC
	2PONbclxl_f2	CGAAGCGCCGGAAGGTACCGAAAGCGAAGCGGTGAAACAGGCGCTGCGTGAAGCGGGTG
	2PONbclxl_f3	CGTTTAGCGATCTGACCAGCCAGCTGCATATTACCCCGGGTACCGCGTATCAGAGC
	2PONbclxl_f4	GATGGTGTGAATTGGGGTCGTATTGTGGCGTTTTTTAGCTTTGGTGGTGCGCTGTGCG
	2PONbclxl_f5	GGTGAGCCGTATTGCGGCGTGGATGGCGACCTATCTGAATGATCATCTGGAACCGTGG
	2PONbclxl_fi1	CGTTTAGCGATCTGACC
	2PONbclxl_r1	CCTTCCGGCGCTTCGGTACGATTTTCTTCCACATCGCTAAACTGGCTCCAGCTATAACC
	2PONbclxl_r2	GGTCAGATCGCTAAACGCACGACGATAACGCAGTTCAAATTCATCACCCGCTTCACGCAG
	2PONbclxl_r3	CCCAATTCACACCATCACGAAACAGTTCATTCACCACCTGTTCAAAGCTCTGATACGCGG
	2PONbclxl_r4	CGCAATACGGCTCACCAGCACCTGCATTTCTTTATCCACGCTTTCCACGCACAGCGCACC
	2PONbclxl_r5	CCATACAGTTCCACAAAGGTATCCCAACCACCATTTTCCTGAATCCACGGTTCCAGATG
	2PONbclxl_ri1	GGTCAGATCGCTAAACG
	2PONbclxl_ri2	CGCAATACGGCTCACC
	pAT223_2PON_f	TTCCGCGGATCCATGAGCCAGAGCAATCGTGAACTGG
	pAT222_2PON_r	TTTGGGAAGCTTCACCATACAGTTCCACAAAGG
	pAT223_2PON_r[a]	TTTGGGAAGCTTCTACTAACCATACAGTTCCACAAAGG
BCL-W	1MK3_bclw_f1	GATACCCGTGCGCTGGTGGCGGATTTTGTGGGTTATAAACTGCGTCAGAAAGGTTATGTG
	1MK3_bclw_f2	CGGATCCGCTGCATCAGGCGATGCGTGCGGCGGGTGATGAATTTGAAACCCGTTTTCGTC
	1MK3_bclw_f2.1	CTGATCCGCTGCATCAGGCGATGC
	1MK3_bclw_f3	GCATGTGACCCCGGGTAGCGCGCAGCAGCGTTTTACCCAGGTGAGCGATGAACTGTTTC
	1MK3_bclw_f4	CGTTTTTTGTGTTTGGTGCGGCGCTGTGCGCGGAAAGCGTGAATAAAGAAATGGAAGTGC
	1MK3_bclw_f5	GGTGGCGTATCTGGAAACCCGTCTGGCGGATTGGATTCATAGCAGCGGTGGTTGGGCGG
	1MK3_bclw_r1	GATGCAGCGGATCCGCCGCCGGACCTTCACCCGGACCCGCACCGCACACATAACCTTTCTG
	1MK3_bclw_r1.2	GCATCGCCTGATGCAGCGGATCAGCCGCCGG
	1MK3_bclw_r2	CCCGGGGTCACATGCAGCTGCGCCGCCAGATCGCTAAAGGTACGACGAAAACGGGTTTC
	1MK3_bclw_r3	CCAAACACAAAAAACGCCACCAGACGACCCCAATTCGGACCACCCTGAAACAGTTCATCGC
	1MK3_bclw_r4	CCAGATACGCCACCATCCATTCCTGCACCTGACCCACCAGCACTTCCATTTCTTTATTC
	1MK3_bclw_r5	CGACGCGCTTCTTCCAGCGCACCATCACCATACAGCGCGGTAAATTCCGCCCAACCACCGC
	1MK3_bclw_fi1	GCATGTGACCCCGGG
	1MK3_bclw_ri1	CCCGGGGTCACATGC
	1MK3_bclw_ri2	CCAGATACGCCACC
	pAT223_1MK3_f	TTCCGCGGATCCGCGACCCCGGCGAGCGCGCCGGATACC
	pAT222_1MK3_r	TTTGGGAAGCTTCACGCACGCTCGCCCAATTACC
	pAT223_1MK3_r[a]	TTTGGGAAGCTTCTACTAACGCACGCTCGCCCAATTACC
MCL-1	2PQK_mcl1_f1	CTGTATCGTCAGAGCCTGGAAATTATTAGCCGTTATCTGCGTGAACAGGCGACCGGTGCG
	2PQK_mcl1_f2	CGACCAGCCGTAAAGCGCTGGAAACCCTGCGTCGTGTGGGTGATGGTGTGCAGCG
	2PQK_mcl1_f3	GCTGCGTAAACTGGATATTAAAAATGAAGATGATGTGAAAAGCCTGAGCCGTGTGATG
	2PQK_mcl1_f4	CCAATTGGGGTCGTATTGTGACCCTGATTAGCTTTGGTGCGTTTGTGGCGAAACATCTG
	2PQK_mcl1_f5	CCGCTGGCGGAAAGCATTACCGATGTGCTGGTGCGTACCAAACGTGATTGGCTGGTG
	2PQK_mcl1_r1	GCTTTACGGCTGGTCGCACCGCTACGACCCATCGGTTTGGTATCTTTCGCACCGGTCGCC
	2PQK_mcl1_r2	CCAGTTTACGCAGCATACCCTGAAACGCGGTTTCATGATTACGCTGCACACCATCACC
	2PQK_mcl1_r3	CAATACGACCCCAATTGGTCACACCATCGCTAAACACATGAATCATCACACGGCTCAGG
	2PQK_mcl1_r4	GCTTTCCGCCAGCGGTTCAATGCAGCTTTCCTGATTAATGGTTTTCAGATGTTTCGCCAC
	2PQK_mcl1_r5	CCACATGAAAAAATTCCACAAAACCATCCCAACCACGCTGTTTCACCAGCCAATCACG
	2PQKmcl1_fi1	GCTGCGTAAACTGG
	2PQK_mcl1_ri1	CCAGTTTACGCAGC
	2PQK_mcl1_ri2	GCTTTCCGCCAGCGG
	pAT223_2PQK_f	TTCCGCGGATCCGATGAACTGTATCGTCAGAGCC
	pAT222_2PQK_r	TTTGGGAAGCTTCATCTTCCACATGAAAAAATTCC
	pAT223_2PQK_r[a]	TTTGGGAAGCTTCTACTAATCTTCCACATGAAAAAATTCC
BIM	hBim_fi1	GCTCTGGTATGCGTCCGGAAATTTGGATTGCGCAGGAACTGCGTCGTATTGGTG
	hBim_ri1	CGCGCATAATACGCATTAAATTCATCACCAATACGACGCAGTTCC
	pAT223_GSGSG_hBim_f	TTCCGCGGATCCGGTTCCGGCTCTGGTATGCGTCCGGAAATTTGG
	pAT223_hBim_r	TTTGGGAAGCTTCTACTAACGCGCATAATACGCATTAAATTC

[a] reverse primer used for the generation of PCR product used to clone into the expression vector pPANK [10,15] expression vector.

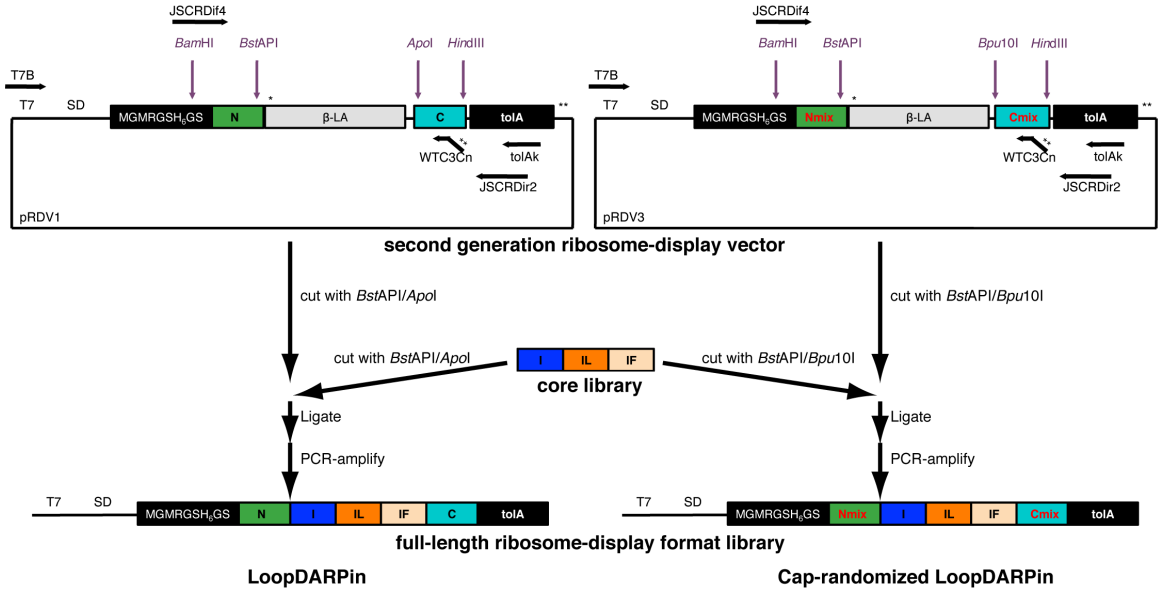
Suppl. Table ST5. Overview of oligonucleotides used for the generation of pAT222-pD and pAT223-pD.

Final vector	Oligonucleotide	Oligonucleotide sequence (5' to 3')
pAT222-pD	pAT222to-pD_f	TTCCGCCGGAATTCATTAAAGAGGAGAAATTAACTATGGCTGGTCTGAACG
	pAT222to-pD_r	TTCCGCGGATCCCATGGAACCTTCGTGCCATTCGATTTTCTGAGCTTCGAAGATATCG
pAT223-pD	pAT223to-pD_f	Phosphate-CATGGGAAGCGGTCATCACCATCACCATCACG
	pAT223to-pD_r	Phosphate-GATCCGTGATGGTGATGGTGATGACCGCTTCC

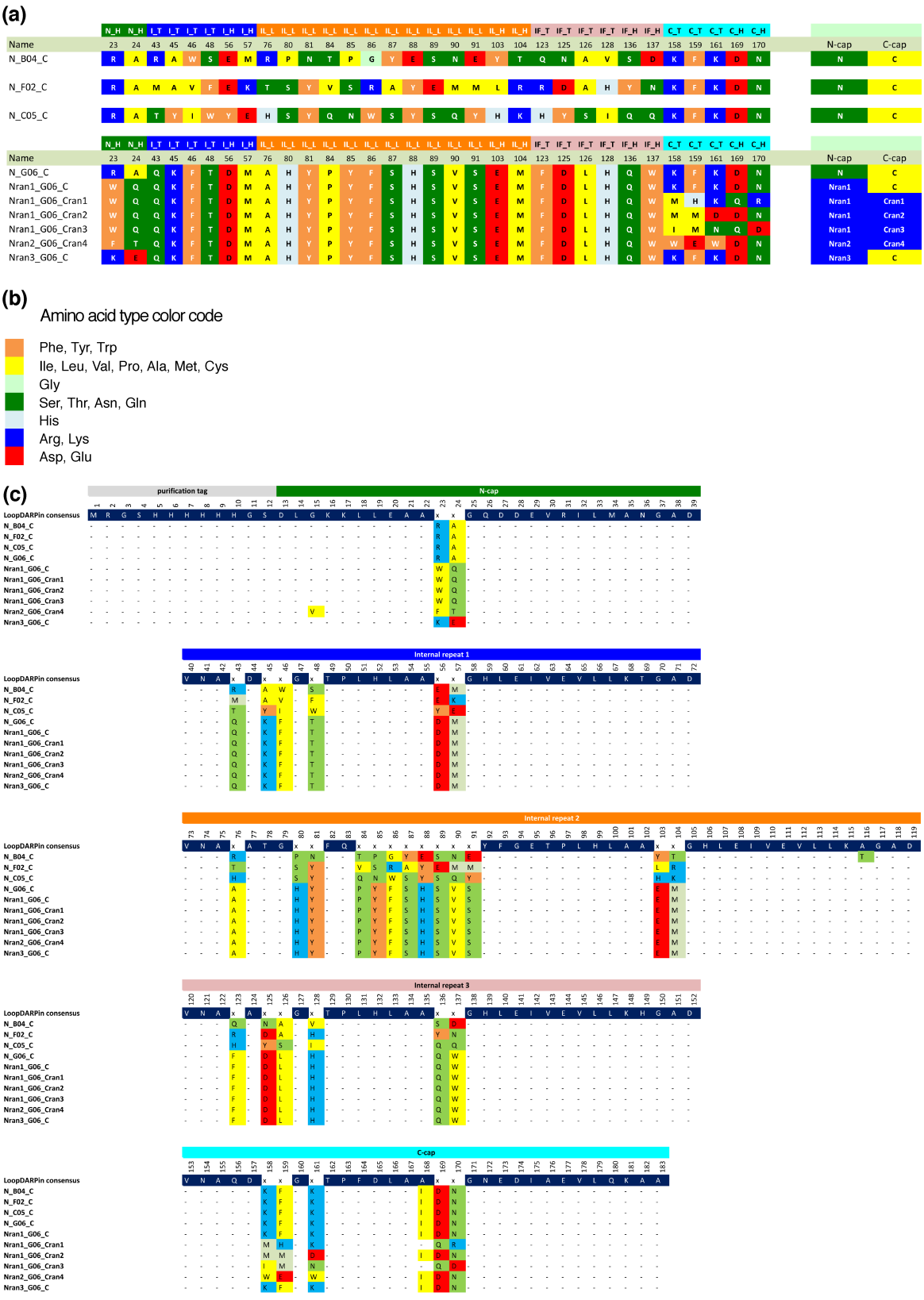
Supplementary Figures



Suppl. Figure S1. Amino acid sequence of single LoopDARPin repeats. Framework positions are colored in blue. Randomized positions are marked with a pink x. Abbreviations given above the randomized positions and numbering given below the sequence can be used to locate the respective position in the LoopDARPin structure (Figure 1 in the main text) and in the amino acid sequence alignments (Suppl. Figure S3 and S9). Nomenclature is as in Figure 1 in the main text and Suppl. Figure S3 and S9. (a) N-terminal capping repeat N or N_{ran}. (b) Internal repeat module I. (c) Internal loop-containing repeat IL. (d) Internal loop-following repeat IF. (e) C-terminal capping repeat C_{old}, C or C_{ran}. Note that C_{old} represents a hybrid between the C-capping repeat from the original DARPin library without loop and the stabilized C-cap version "C", as explained in the main text. The differences between C and C_{old} are in positions 11, 14, and 15.

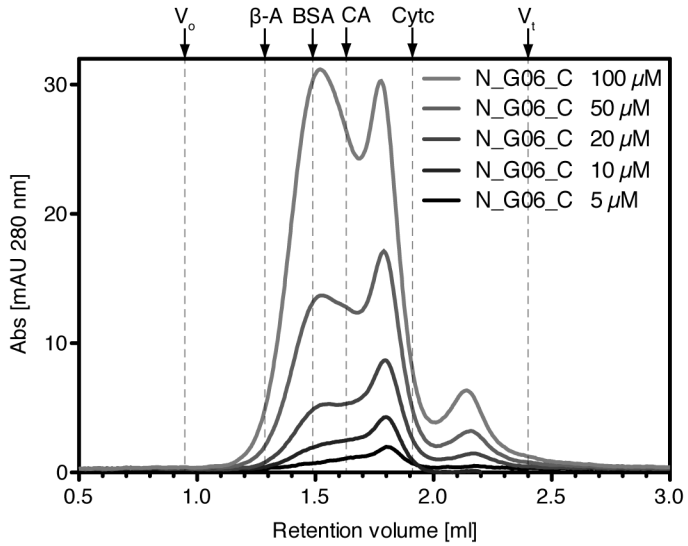


Suppl. Figure S2. Our new assembly strategy of full-length ribosome-display format LoopDARPin libraries is shown. The introduction of restriction sites present at repeat boundaries allows the simple fusion of capping repeats together with essential flanking sequence features present in the pRDV ribosome-display vector, such as the promoter and Shine-Dalgarno sequence upstream and the tether and RNA hairpin downstream [11,12] to the I-IL-IF core library. Either unrandomized (sequence-defined) or randomized capping repeats may be added by using our two newly generated second generation pRDV ribosome-display vectors pRDV1 or pRDV3, respectively. Primers used during ribosome-display selections and screening are indicated by arrows in pRDV1 and pRDV3.

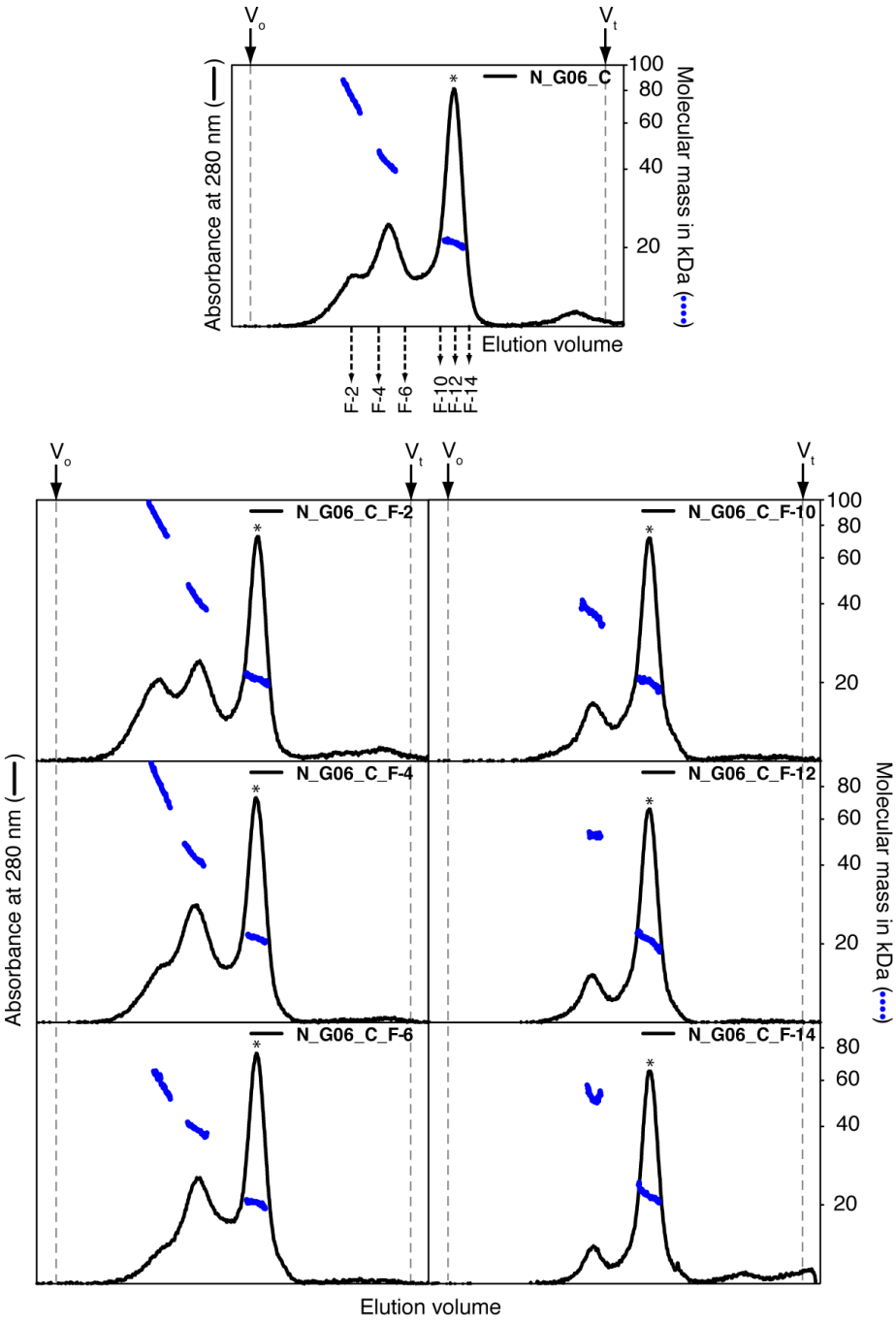


Suppl. Figure S3. (a) Relevant sequence positions of unselected LoopDARPins characterized in detail. Only randomized positions (or nonrandomized equivalents) are shown in single letter code and are colored according to amino acid type. Randomized positions are identified by a number, corresponding to the amino acid position in the full-

length expressed sequence including purification tags, and by a string of the following composition "repeat"_"secondary structure element", which describes the location of the position in the structure. The ending _H denotes randomized positions in helical parts, and _T in the β-turns. Randomized positions of the loop are marked with IL_L. These identifiers are colored as follows: N-terminal capping repeat (N-cap, N or N_{ran}, green background), the three internal repeat modules (I = internal repeat, blue background; IL = loop-containing internal repeat, orange background and IF = loop-following internal repeat, pink background) and the C-terminal capping repeat (C-cap, C or C_{ran}, cyan background). At the right, the nature of the capping repeat is indicated again for clarification. (b) Amino acid type coloring code. (c) Full-length sequences of LoopDARPins shown in (a). The LoopDARPin consensus sequence is shown in the first line. Randomized positions in the LoopDARPin consensus are indicated by X. In the lower lines, only residues that differ from the consensus sequence are printed. Sequences are shown in single letter code. Sequence positions are identified by a number corresponding to the amino acid position in the full-length expressed sequence. Coloring as in (a).

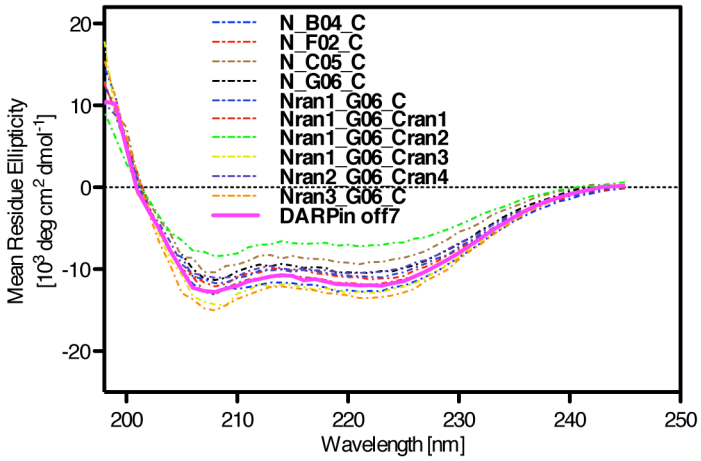


Suppl. Figure S4. Analytical size-exclusion chromatography with increasing concentrations of LoopDARPin N_G06_C. The mass of the monomer calculated from the sequence is 19.7 kDa. With increasing concentrations, the fraction of oligomeric species increases, illustrating that the observed oligomerization is concentration-dependent. The void volume (V_o = 0.95 ml), the total volume (V_t = 2.4 ml) and the molecular mass standards (β-amylase with an apparent mass of 200 kDa, BSA with an apparent mass of 66 kDa, carbonic anhydrase with an apparent mass of 29 kDa and cytochrome c with an apparent mass of 12.4 kDa) are indicated by broken gray lines in the graph. The small peak at 2.15 ml was identified to be DNA or RNA.

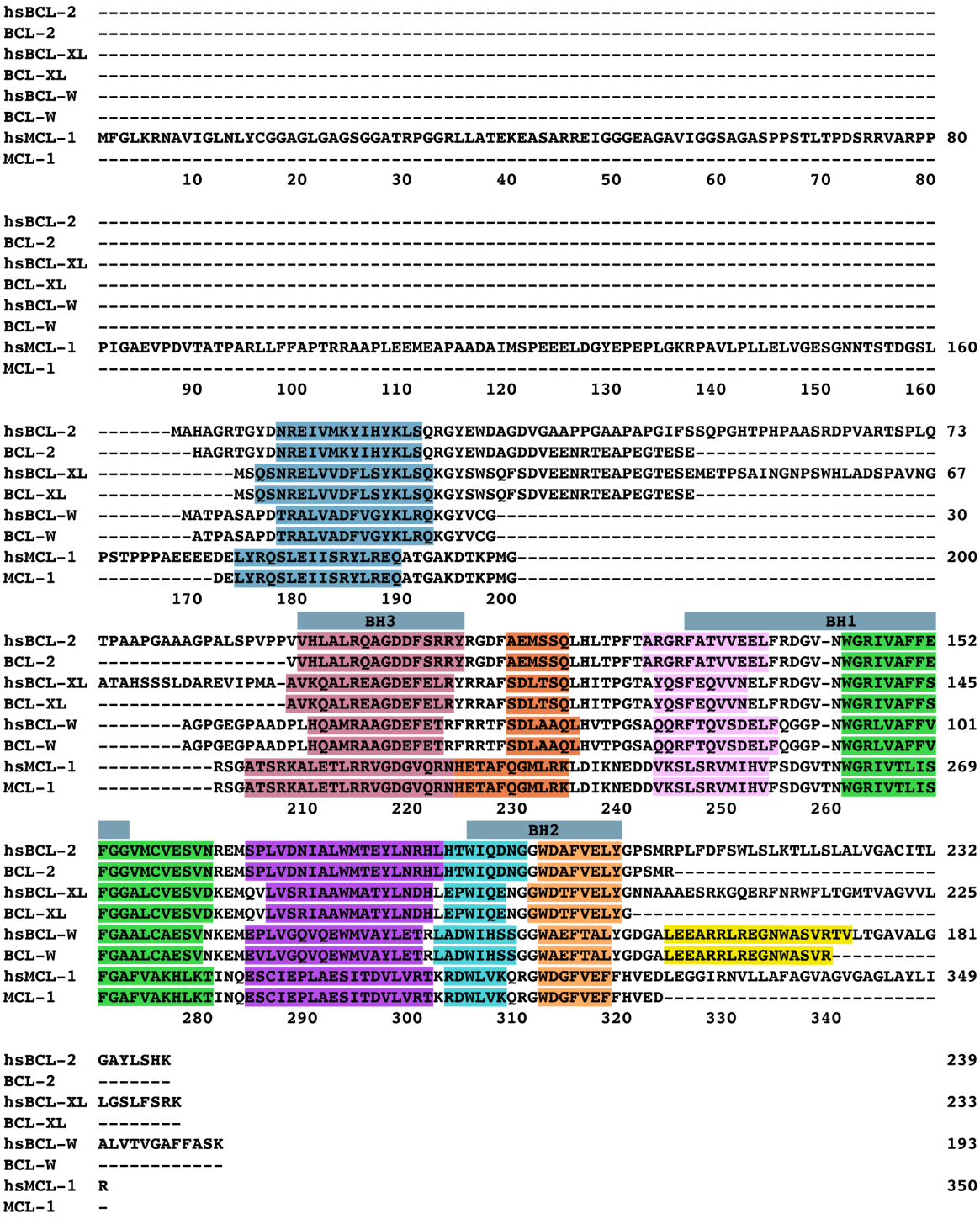


Suppl. Figure S5. Reversibility of oligomerization tested by SEC-MALS. As in Figure 4 in the main text, the chromatograms show the absorption at 280 nm (left-hand y-axis). Indicated by a line of blue dots across the eluting peaks, the right-hand y-axis shows the molecular mass of the eluting particles as determined by MALS. Approximate calculated molecular weights are: monomer 20 kDa, dimer 40 kDa, trimer 60 kDa and tetramer 80 kDa. For exact calculated molecular weights of the monomer see Table 1. Peaks corresponding to the monomeric species are denoted with an asterisk (*). The void volume (V_o = 10 ml) and the total volume (V_t = 24 ml) are indicated by broken gray lines in the graph. Re-injection of either collected monomer or oligomer fractions (shown with

arrows below top chromatogram) results in a distribution close to the original chromatogram, showing that oligomers can be reversibly transformed back into monomers, and vice versa. The fraction injected is indicated in the lower chromatograms. Along with additional results (see text) this suggests that the observed oligomer formation is the reversible association between folded monomers. The original chromatogram was recorded at a concentration of 1 mg/ml, whereas fractions F-2 to F-6 and fractions F-10 to F-14 were re-injected at a concentration of 0.5 and 0.25 mg/ml, respectively. A reduced oligomer formation after re-injection of fractions F-10 to F-14 might therefore originate from a lower concentration of the re-injected sample.

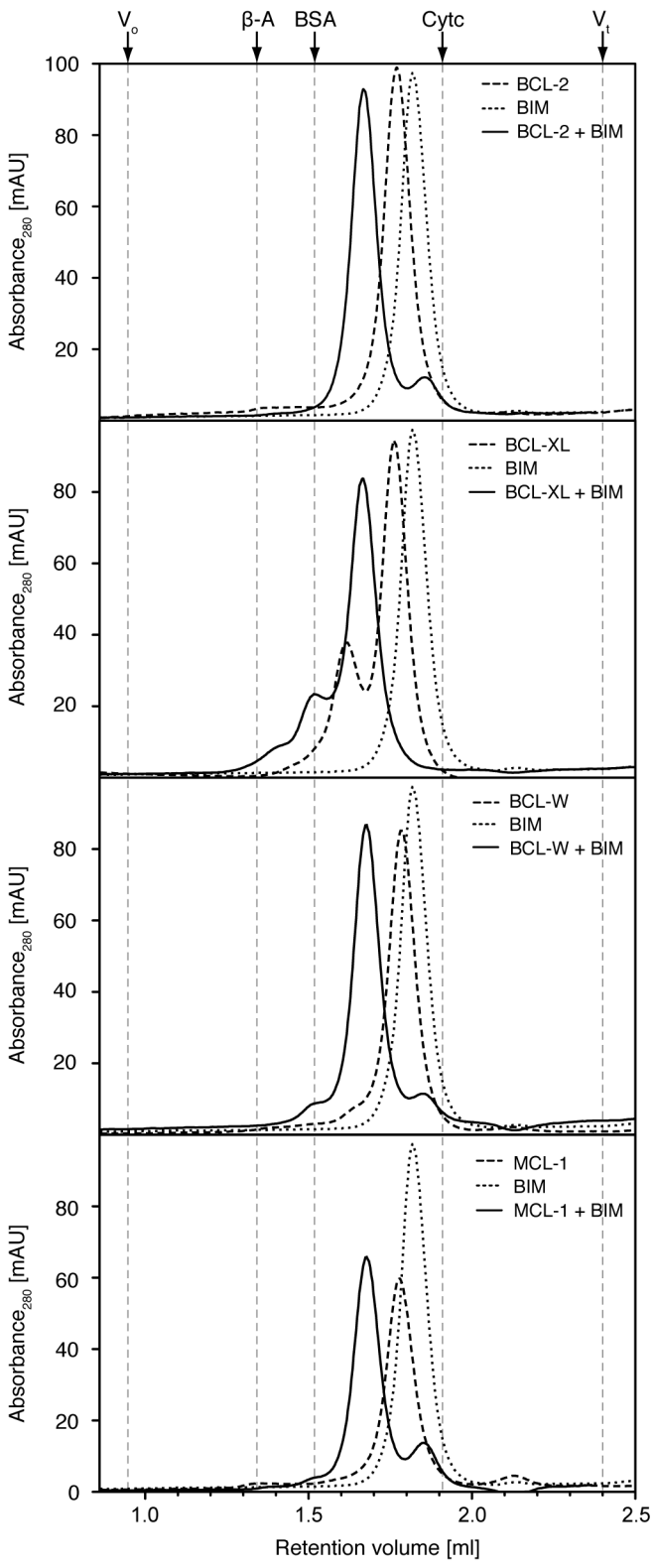


Suppl. Figure S6. Circular dichroism spectra of the 10 investigated randomly chosen unselected LoopDARPin library members and the conventional DARPin off7 [21] are shown. All LoopDARPins exhibit spectra and α-helical content very similar to the spectrum of the conventional DARPin off7 and other conventional DARPins [10]. Only LoopDARPin $N_{ran}1\_G06\_C_{ran}2$ deviates slightly from the other examined proteins and shows reduced α-helical signal compared to all other proteins tested.



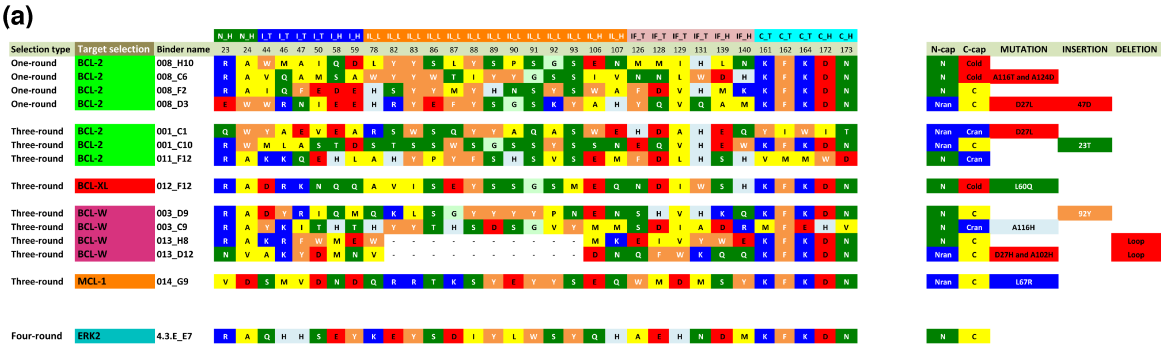
Suppl. Figure S7. Structure-based sequence alignment of the anti-apoptotic BCL-2 family members used in the present study. Sequences of the full-length homo sapiens (hs) protein are included in the alignment and are listed above the final construct used for each family member. The definition of N- and C-terminal boundaries in the finally used sequences was guided by the resolved residues in the structures used for structure-based sequence alignment: BCL-2 (PDB-ID: 1G5M) [23], BCL-XL (PDB-ID: 2PON) [24], BCL-W (PDB-ID: 1MK3) [25,26] and MCL-1 (PDB-ID: 2PQK) [27,28]. For structures, see Figure 6 in

the main text. The BH domains are indicated by colored bars above the sequence and helix positions as colored bars on the sequence. The displaceable [29,30] groove-binding helix α8 of BCL-W is highlighted in yellow. As in ref. [23] a truncated modified α1-α2 loop version of BCL-2 was used.

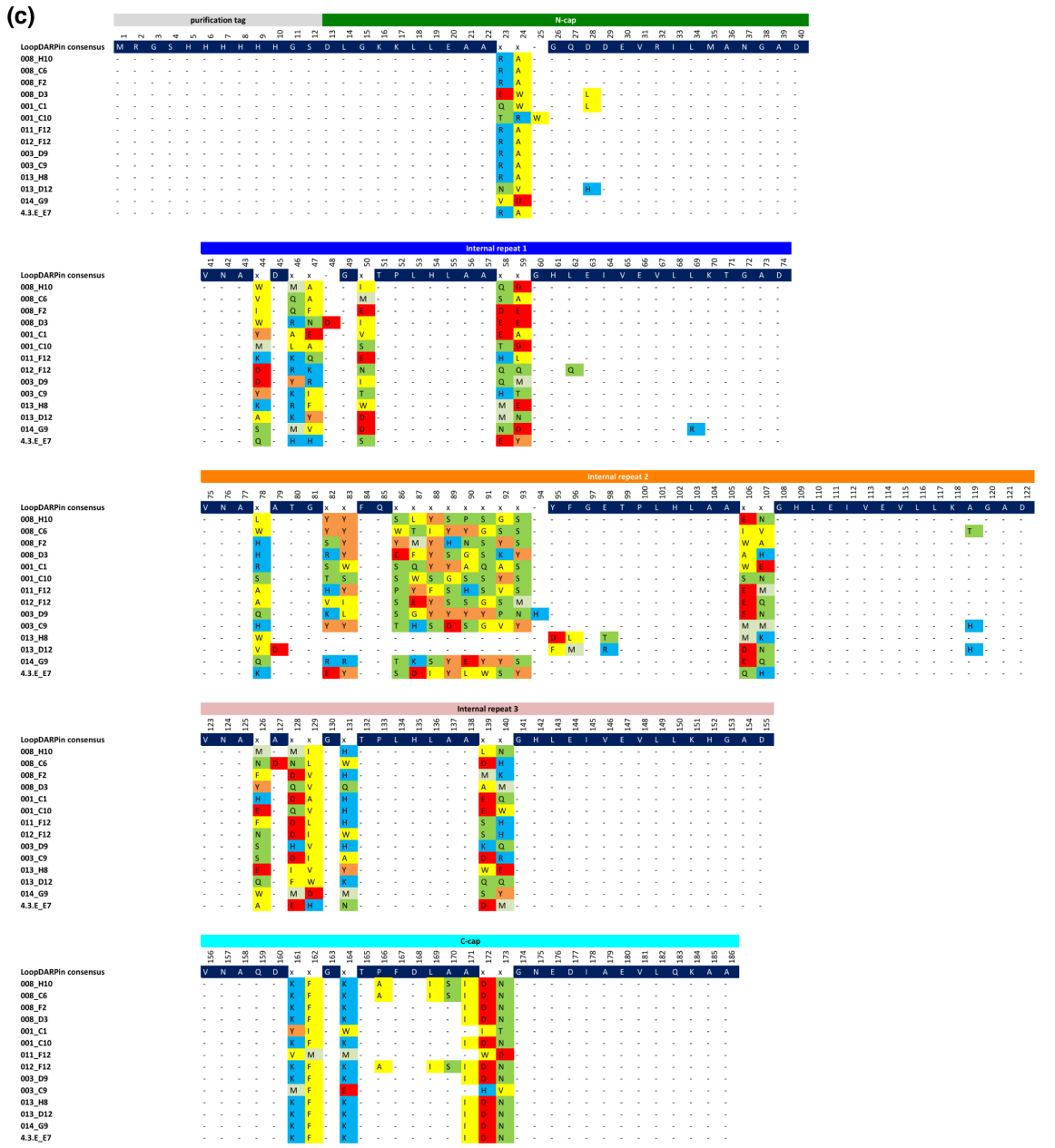


Suppl. Figure S8. Test of functional integrity of BCL-2 family members. The correct folding of anti-apoptotic BCL-2 family members upon *E. coli* expression was analyzed through complex formation with a pD fusion of their natural BH3-only BIM peptide ligand, detected by size-exclusion chromatography. The BIM peptide binds to a surface-

exposed hydrophobic groove, which is essential for the proper function of anti-apoptotic BCL-2 family members. Superposition of analytical SEC elution profiles of target proteins (approximate calculated molecular mass of 21 kDa), pD-BIM peptide ligand fused to pD with a calculated molecular mass of 16 kDa and complexes thereof are shown. Free target proteins (dashed line), free pD-BIM peptide ligand (short dashed line) and target proteins complexed with the pD-BIM peptide ligand (solid line) are shown. Protein elution was followed by UV absorption at 280 nm. Free proteins and ligands as well as equimolar complexes thereof were used at a concentration of 10 μM. For all four target proteins BCL-2, BCL-XL, BCL-W and MCL-1 a clear shift of the complex peak (solid line) to a higher molecular weight (smaller retention volume) can be observed. In all cases except BCL-XL, minor tailing shoulders at higher retention volumes are present. These shoulders can be attributed to either uncomplexed target protein or uncomplexed ligand. Importantly, the observed quantitative target/pD-BIM complex formation supports the correct folding of the used BCL-2 family members. The void volume (V_o = 0.95 ml), the total volume (V_t = 2.4 ml) and the molecular mass standards (β-amylase with an apparent mass of 200 kDa, BSA with an apparent mass of 66 kDa and cytochrome c with an apparent mass of 12.4 kDa) are indicated by broken gray lines in the graph.

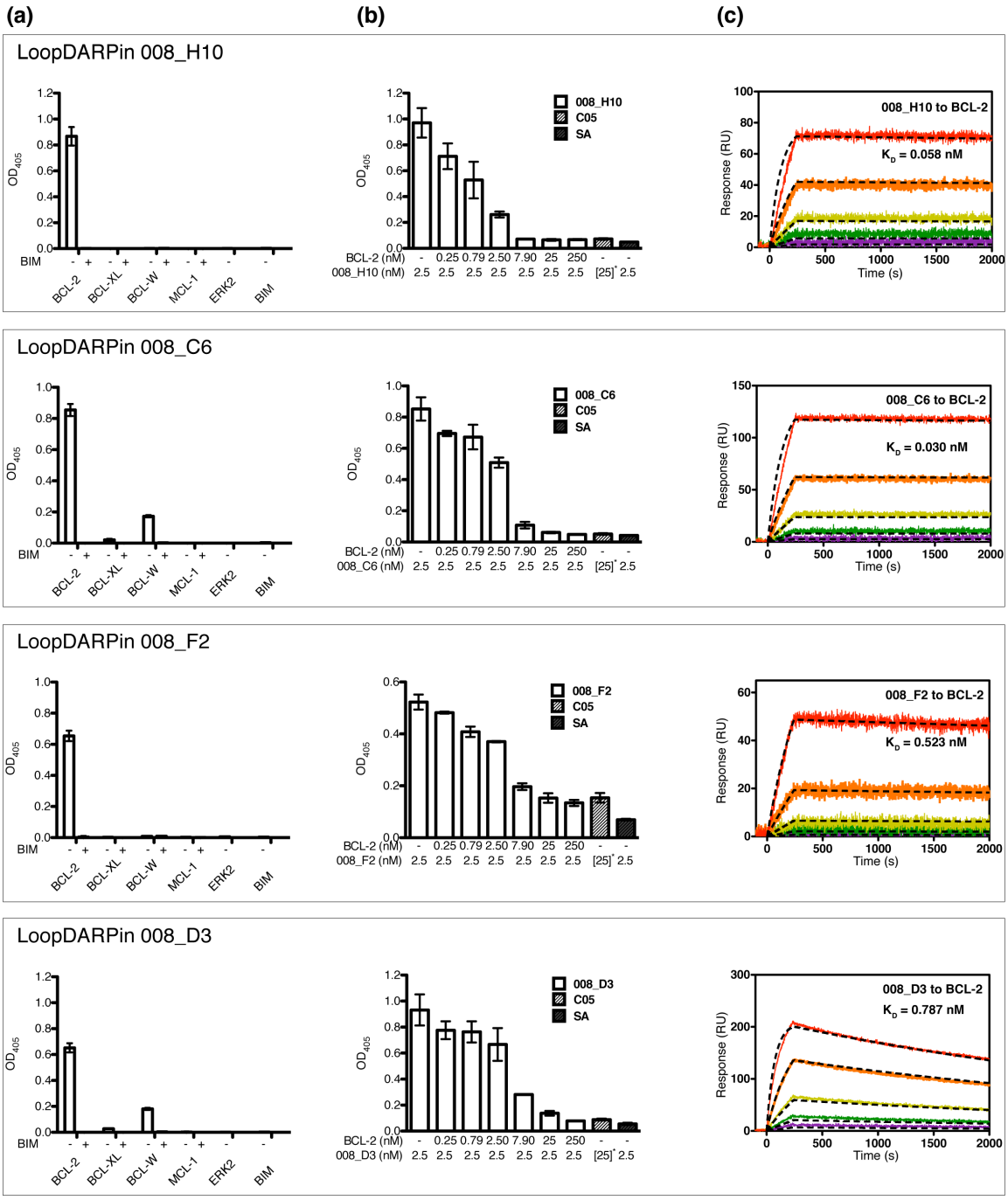

(a)
Selection type
Target selection
Binder name
One-round
BCL-2
008_H10
008_C6
008_F2
008_D3
Three-round
001_C1
001_C10
011_F12
BCL-XL
012_F12
BCL-W
003_D9
003_C9
013_H8
013_D12
MCL-1
014_G9
Four-round
ERK2
4.3.E_E7
N-cap
C-cap
MUTATION
INSERTION
DELETION
A110T and A124D
D27L
47D
23T
92Y
A116H
Loop
D27H and A102H
L67R
L60Q

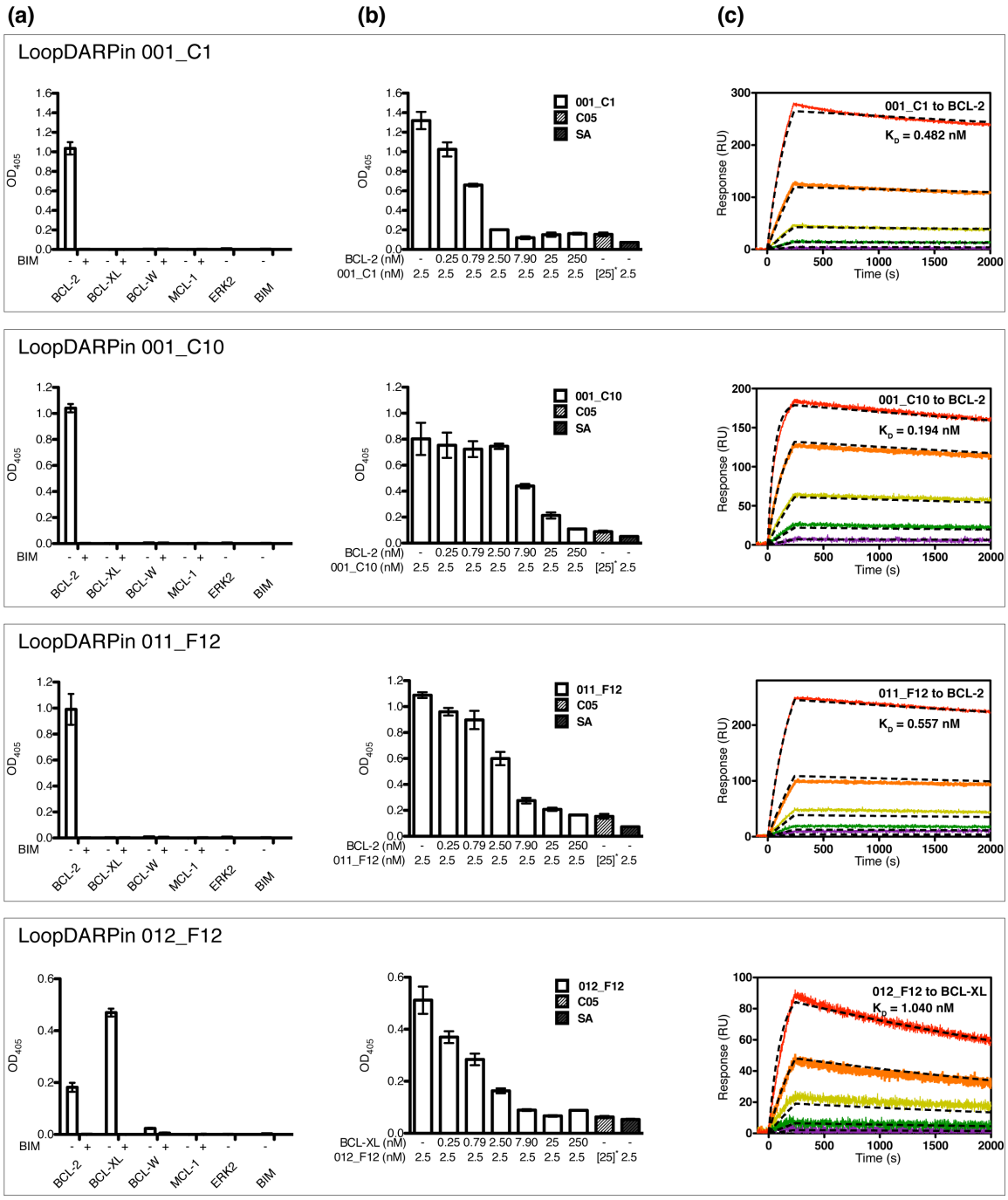

(b)
Amino acid type color code
Phe, Tyr, Trp
Ile, Leu, Val, Pro, Ala, Met, Cys
Gly
Ser, Thr, Asn, Gln
His
Arg, Lys
Asp, Glu

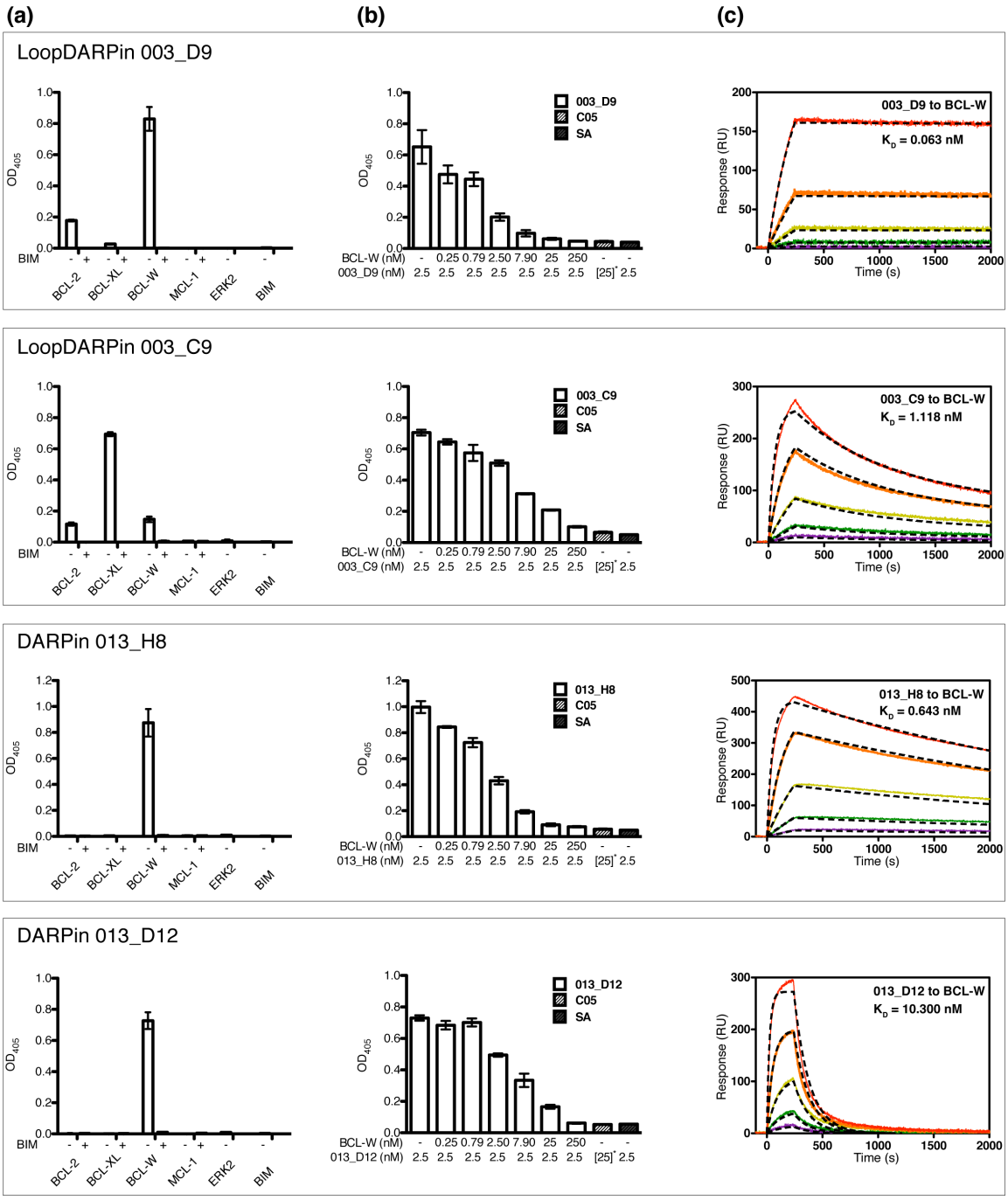

(c)
purification tag
N-cap
Internal repeat 1
Internal repeat 2
Internal repeat 3
C-cap
LoopDARPin consensus
008_H10
008_C6
008_F2
008_D3
001_C1
001_C10
011_F12
012_F12
003_D9
003_C9
013_H8
013_D12
014_G9
4.3.E_E7

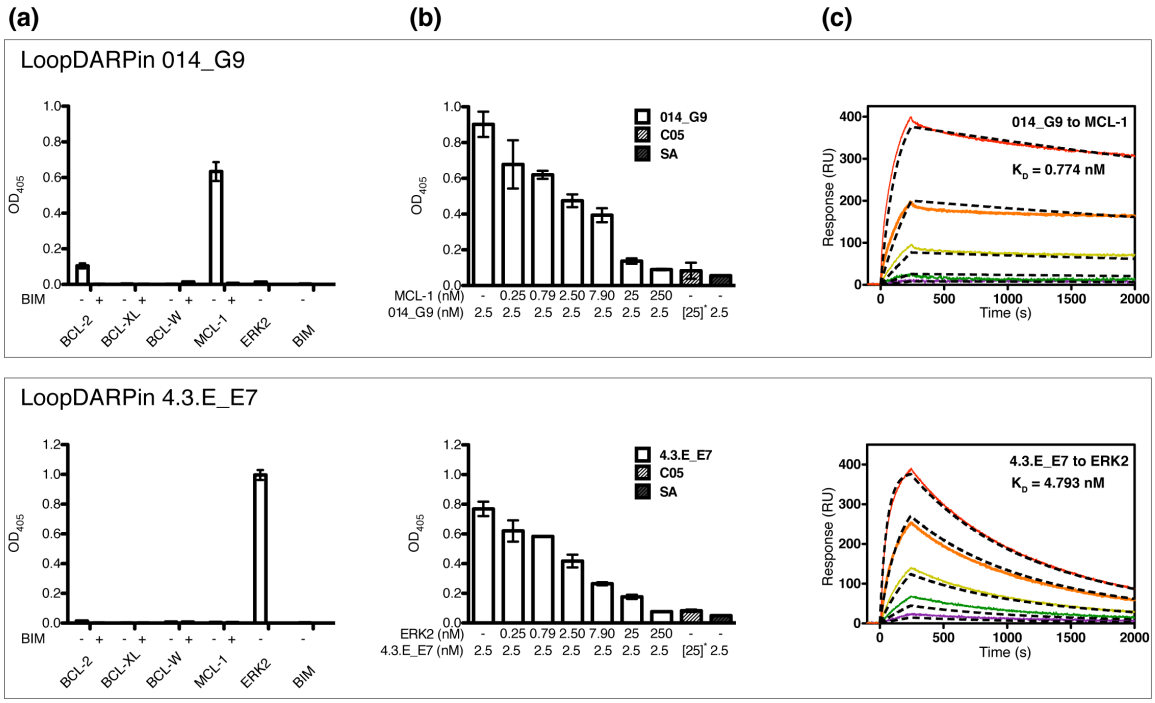
Suppl. Figure S9. (a) Relevant sequence positions of all binders analyzed in detail. Only randomized positions are shown in single letter code and are colored according to amino acid type. Randomized positions are identified by a number, corresponding to the amino acid position in the full length expressed sequence including purification tags, and by a string of the composition "repeat"_"secondary structure element", which describes the location of the position in the structure. The ending _H denotes randomized positions in helical parts, and _T in the β-turns. Randomized positions of the loop are marked with IL_L. These identifiers are colored as follows: N-terminal capping repeat (N-cap, N or N_{ran}, green background), the three internal repeat modules (I = internal repeat, blue background; IL = loop-containing internal repeat, orange background and IF = loop-following internal repeat, pink background) and the C-terminal capping repeat (C-cap, C or C_{ran}, cyan background). At the right, the nature of the capping repeat is indicated again for clarification. In addition to that, mutations, insertions or deletions are listed. (b) Amino acid type coloring code. (c) Full-length sequences of LoopDARPins shown in (a). The LoopDARPin consensus sequence is shown in the first line. Randomized positions in the LoopDARPin consensus are indicated by X. In the lower lines, only residues that differ from the consensus sequence are printed. Sequences are shown in single letter code. Sequence positions are identified by a number corresponding to the amino acid position in the full-length expressed sequence. Coloring as in (a).

As intended from the biased design of the naïve library, randomized loop positions are rich in Tyr, Ser and Gly. However, all other possible amino acids (except Cys) are present as well, supporting the chosen randomization strategy. For the BCL-2 family binders shared sequence patters are present for binders obtained from the same target selection, but also among binders from different target selections. Among several binders, a basic residue at position 46, a cluster of negatively charged residues in the first internal repeat, His at position 78, 131 and 140, Ser or Gly at position 91 and Ser at position 93 or residues with aliphatic side chains at position 129 appear. Nevertheless, it is difficult to conclude specificity patterns from these sequence patterns, as binders, which share the mentioned residues still have different specificities, or binders with different randomized residues bind to the same target. For example, both 008_H10 (BCL-2 selection), which exclusively binds BCL-2 and 003_D9 (BCL-W selection), which binds BCL-2, BCL-XL and BCL-W possess an aliphatic side chains at position 129 and a His at position 131. 008_C6 and 008_F2, both from the BCL-2 selection possess rather different randomized positions, but still bind the same target with high affinity.


(a)
(b)
(c)
LoopDARPin 008_H10
OD405
BIM
BCL-2
BCL-XL
BCL-W
MCL-1
ERK2
BIM
008_H10
C05
SA
BCL-2 (nM) - 0.25 0.79 2.50 7.90 25 250 - -
008_H10 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
008_H10 to BCL-2
K_D = 0.058 nM
Response (RU)
Time (s)
LoopDARPin 008_C6
008_C6
C05
SA
BCL-2 (nM) - 0.25 0.79 2.50 7.90 25 250 - -
008_C6 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
008_C6 to BCL-2
K_D = 0.030 nM
Response (RU)
Time (s)
LoopDARPin 008_F2
008_F2
C05
SA
BCL-2 (nM) - 0.25 0.79 2.50 7.90 25 250 - -
008_F2 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
008_F2 to BCL-2
K_D = 0.523 nM
Response (RU)
Time (s)
LoopDARPin 008_D3
008_D3
C05
SA
BCL-2 (nM) - 0.25 0.79 2.50 7.90 25 250 - -
008_D3 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
008_D3 to BCL-2
K_D = 0.787 nM
Response (RU)
Time (s)

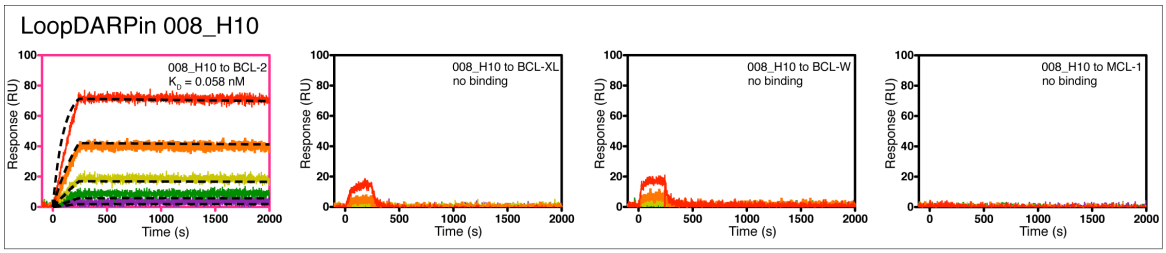

(a)
(b)
(c)
LoopDARPin 001_C1
OD405
BIM
BCL-2
BCL-XL
BCL-W
MCL-1
ERK2
BIM
BCL-2 (nM)
001_C1 (nM)
001_C1
C05
SA
001_C1 to BCL-2
KD = 0.482 nM
Response (RU)
Time (s)
LoopDARPin 001_C10
001_C10
001_C10 to BCL-2
KD = 0.194 nM
LoopDARPin 011_F12
011_F12
011_F12 to BCL-2
KD = 0.557 nM
LoopDARPin 012_F12
012_F12
012_F12 to BCL-XL
KD = 1.040 nM
BCL-XL (nM)
012_F12 (nM)

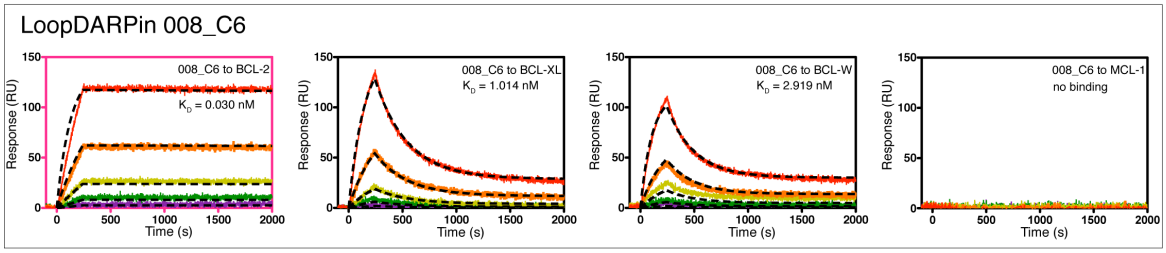

(a)
(b)
(c)
LoopDARPin 003_D9
OD405
BIM
BCL-2
BCL-XL
BCL-W
MCL-1
ERK2
BIM
BCL-W (nM) - 0.25 0.79 2.50 7.90 25 250 - -
003_D9 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
003_D9
C05
SA
003_D9 to BCL-W
K_D = 0.063 nM
Response (RU)
Time (s)
LoopDARPin 003_C9
003_C9 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
003_C9
C05
SA
003_C9 to BCL-W
K_D = 1.118 nM
DARPin 013_H8
013_H8 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
013_H8
C05
SA
013_H8 to BCL-W
K_D = 0.643 nM
DARPin 013_D12
013_D12 (nM) 2.5 2.5 2.5 2.5 2.5 2.5 2.5 [25]* 2.5
013_D12
C05
SA
013_D12 to BCL-W
K_D = 10.300 nM

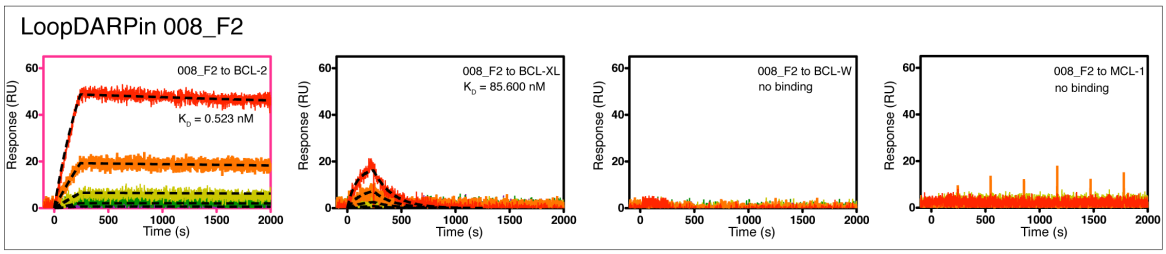


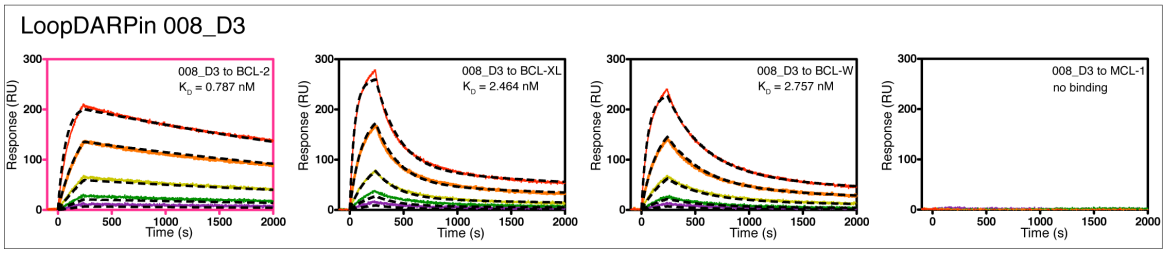
Suppl. Figure S10. Specificity and affinity results of all selected binders for the target they originally were selected against. (a) Background-corrected ELISA signals to determine binder specificity. The interaction of all selected LoopDARPins (5 nM) with immobilized anti-apoptotic BCL-2 family members (–BIM) can be blocked by preincubation with a stoichiometric excess of the natural BIM peptide ligand (+BIM), indicating that the groove-directing pre-panning ribosome-display strategy worked and that selected BCL-2 family binders bind into or near the surface-exposed groove. Direct interaction of binders with the BIM peptide does not occur (last column). Note that the background binding of LoopDARPins/DARPins to the surface without immobilized target (< 10% of total binding) has been subtracted. (b) Competition ELISA illustrating the interaction between binders and their original target. Binders (final concentration: 2.5 nM) were incubated with varying concentrations of free target before binding on immobilized target. Binding of all selected binders can be specifically inhibited by increasing concentrations of free target in solution. The unselected LoopDARPin C05 (hatched column, denoted with asterisk (*)) showed no interaction with any target at 25 nM, giving a signal identical to that of the control (SA [streptavidin only]: 2.5 nM selected binder without immobilized target). (c) Surface plasmon resonance (SPR) analyses. Different concentrations of LoopDARPin (50, 15.81, 5, 1.58, 0.5 and 0 nM; red to gray) were injected simultaneously on parallel lanes during one run, followed by washing with buffer flow. The global fit is indicated by black dashed lines (see Suppl. Table 2 for the

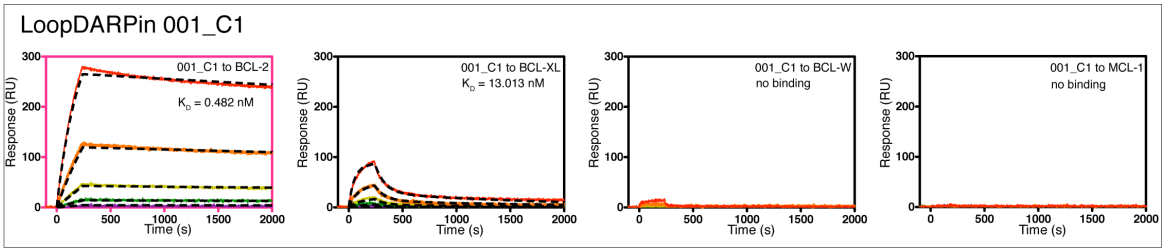
extracted kinetic data and Figure 8 in the main text and Suppl. Results for the used fitting model).

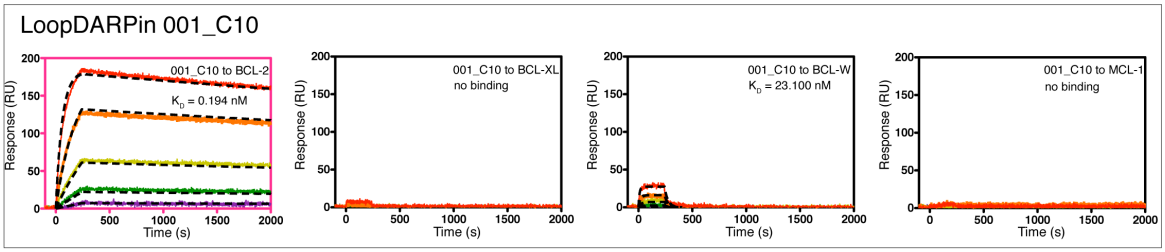

LoopDARPin 008_H10
008_H10 to BCL-2
K_D = 0.058 nM
008_H10 to BCL-XL
no binding
008_H10 to BCL-W
no binding
008_H10 to MCL-1
no binding
Response (RU)
Time (s)

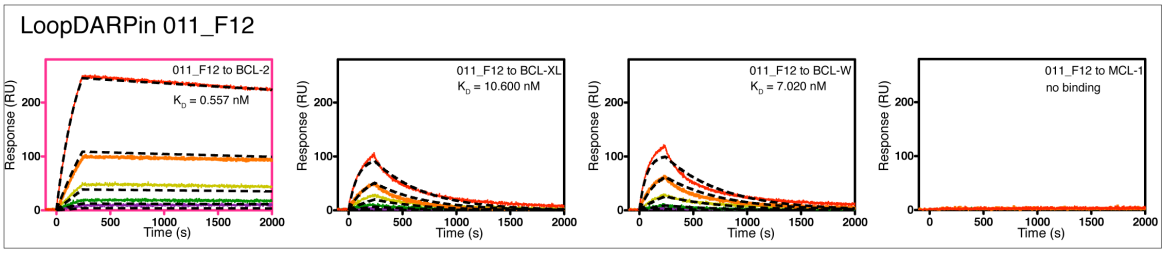

LoopDARPin 008_C6
008_C6 to BCL-2
K_D = 0.030 nM
008_C6 to BCL-XL
K_D = 1.014 nM
008_C6 to BCL-W
K_D = 2.919 nM
008_C6 to MCL-1
no binding
Response (RU)
Time (s)

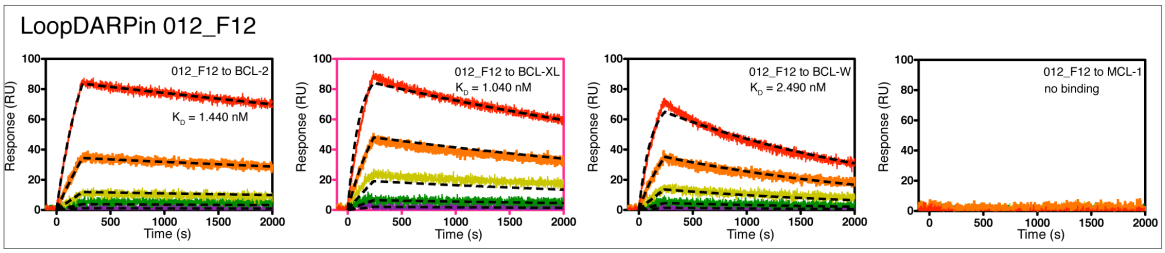

LoopDARPin 008_F2
008_F2 to BCL-2
K_D = 0.523 nM
008_F2 to BCL-XL
K_D = 85.600 nM
008_F2 to BCL-W
no binding
008_F2 to MCL-1
no binding
Response (RU)
Time (s)

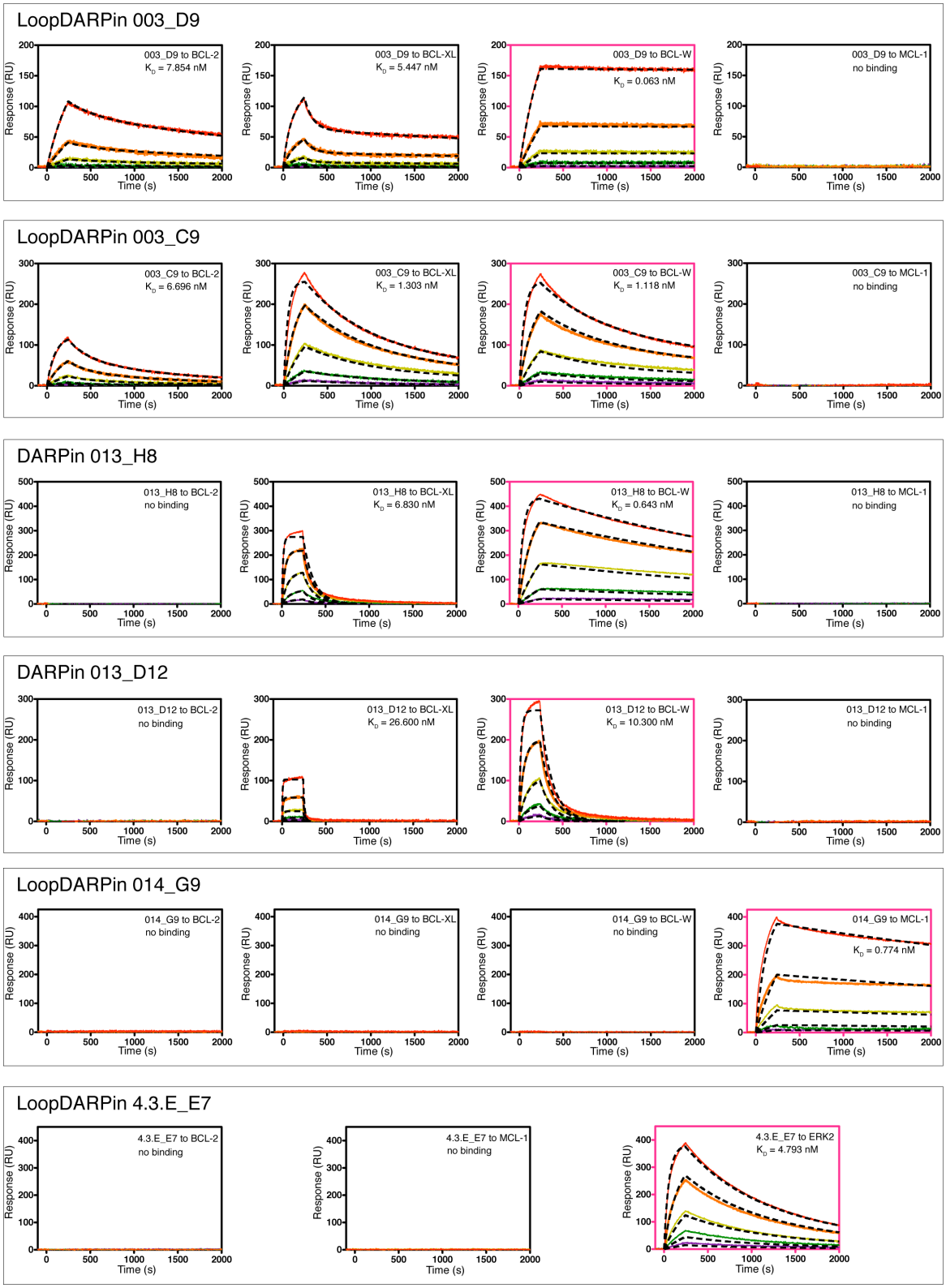

LoopDARPin 008_D3
008_D3 to BCL-2
K_D = 0.787 nM
008_D3 to BCL-XL
K_D = 2.464 nM
008_D3 to BCL-W
K_D = 2.757 nM
008_D3 to MCL-1
no binding
Response (RU)
Time (s)


LoopDARPin 001_C1
001_C1 to BCL-2
K_D = 0.482 nM
001_C1 to BCL-XL
K_D = 13.013 nM
001_C1 to BCL-W
no binding
001_C1 to MCL-1
no binding
Response (RU)
Time (s)


LoopDARPin 001_C10
001_C10 to BCL-2
K_D = 0.194 nM
001_C10 to BCL-XL
no binding
001_C10 to BCL-W
K_D = 23.100 nM
001_C10 to MCL-1
no binding
Response (RU)
Time (s)

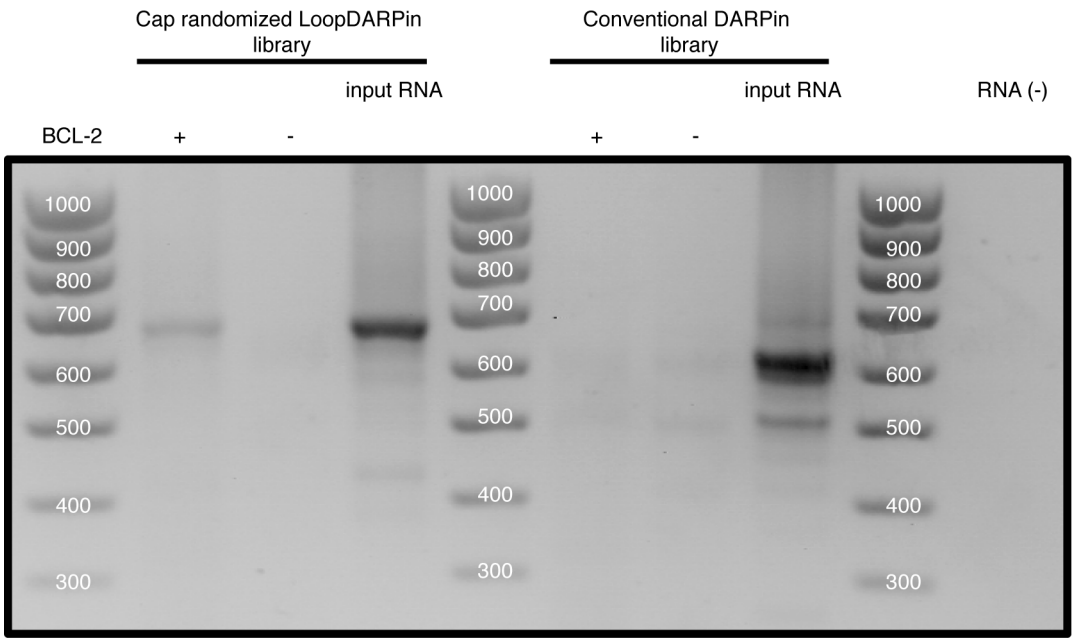

LoopDARPin 011_F12
011_F12 to BCL-2
K_D = 0.557 nM
011_F12 to BCL-XL
K_D = 10.600 nM
011_F12 to BCL-W
K_D = 7.020 nM
011_F12 to MCL-1
no binding
Response (RU)
Time (s)


LoopDARPin 012_F12
012_F12 to BCL-2
K_D = 1.440 nM
012_F12 to BCL-XL
K_D = 1.040 nM
012_F12 to BCL-W
K_D = 2.490 nM
012_F12 to MCL-1
no binding
Response (RU)
Time (s)

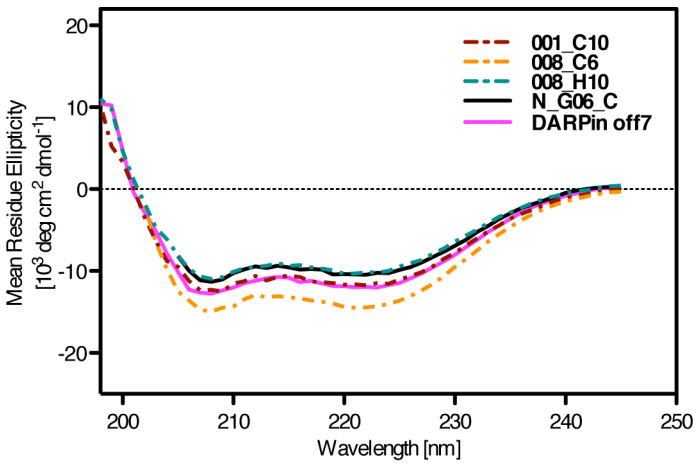


Suppl. Figure S11. Specificity and affinity results for all selected binders determined by SPR measurements. The order of the presented sensorgrams is according to the listed K_Ds in Figure 8 in the main text and Suppl. Table S2. Thus, for each binder, the interactions with BCL-2, BCL-XL, BCL-W and MCL-1 are shown from left to right. Sensorgrams, which show the interaction of binders with the cognate target (that they have originally

been selected against), are highlighted by a pink frame, as in Figure 8 in the main text. Note that for binders, which interact with more than one BCL-2 family member, the interaction with the original target is always the strongest. Different concentrations of LoopDARPin/DARPin (50, 15.81, 5, 1.58, 0.5 and 0 nM; red to gray) were injected simultaneously onto different lanes during one run, followed by washing with buffer flow. The global fit is indicated by black dashed lines (see Suppl. Table ST2 for the extracted kinetic data and Figure 8 in the main text and Suppl. Results for the used fitting model).



Suppl. Figure S12. Recovery over background of reverse-transcribed and PCR-amplified RNA of the one-round ribosome-display (RD) selection. The outcome of RD was monitored at the level of RT-PCR product yield by agarose gel electrophoresis. The RT-PCR yields originating from selections containing target (+) was compared to the yields from selections containing only BSA-blocked streptavidin-coated magnetic beads, but no target (–). The amount and quality of the selection input RNA was verified separately for each selection by RT-PCR amplification (input RNA). Strong recovery over background (quantitative difference between target (+) and no-target lane (-)) indicated the successful selection of binders from the naïve cap-randomized LoopDARPin library within only one round of RD. Expected sizes of RT-PCR reaction products are 654 bp (LoopDARPins) and 612 bp (conventional DARPins without loop).



Suppl. Figure S13. Circular dichroism spectra of the three investigated selected LoopDARPins, the unselected LoopDARPin N_G06_C and the conventional DARPin off7 [21,31] are shown. All LoopDARPins exhibit spectra and α-helical content very similar to the spectra of unselected LoopDARPins exemplified by N_G06_C and conventional DARPins exemplified by DARPin off7.

Initial hits

						Cap combinations						
	N	**N_{ran}**	**C_{old}**	**C**	**C_{ran}**	**N-C_{old}**	**N-C**	**N-C_{ran}**	**N_{ran}-C_{old}**	**N_{ran}-C**	**N_{ran}-C_{ran}**	**total**
observed	14	12	10	10	6	8	4	2	2	6	4	26
observed in %	54	46	38	38	23	31	15	8	8	23	15	100
expected in %	67	33	25	25	50	17	17	33	8	8	17	100

Final hits

						Cap combinations						
	N	**N_{ran}**	**C_{old}**	**C**	**C_{ran}**	**N-C_{old}**	**N-C**	**N-C_{ran}**	**N_{ran}-C_{old}**	**N_{ran}-C**	**N_{ran}-C_{ran}**	**total**
observed	8	5	3	7	3	3	3	2	0	4	1	13
observed in %	62	38	23	54	23	23	23	15	0	31	8	100
expected in %	67	33	25	25	50	17	17	33	8	8	17	100

Suppl. Figure S14. Observed occurrence of certain cap variants and combinations of N- and C-cap variants in both initial and final hit binders in comparison to the expected occurrence derived from the composition of the naïve library.