jmb Article

Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High Specificity

Gautham Varadamsetty, Dirk Tremmel, Simon Hansen, Fabio Parmeggiani and Andreas Plückthun

Department of Biochemistry, *University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

Correspondence to Andreas Plückthun: *plueckthun@bioc.uzh.ch*
http://dx.doi.org/10.1016/j.jmb.2012.08.029
Edited by S. Koide

Abstract

Designed Armadillo repeat proteins (ArmRPs) are a novel class of binding proteins intended for general modular peptide binding and have very favorable expression and stability properties. Using a combination of sequence and structural consensus analyses, we generated a 42-amino-acid designed Armadillo repeat module with six randomized positions, having a theoretical diversity of 9.9×10^6 per repeat. Structural considerations were used to replace cysteine residues, to define less conserved positions and to decide where to introduce randomized amino acid residues for potential interactions with the target peptide. Based on these concepts, combinatorial libraries of designed ArmRPs were assembled. The most stable version of designed ArmRP in library format was the N5C format, with three randomized library repeat modules flanked by full consensus repeat modules on either side and, in turn, flanked by N- and C-terminal capping repeats. Unselected members of this library were well expressed in the *Escherichia coli* cytoplasm, monomeric and showed the expected CD spectra and cooperative unfolding. N5C libraries were used in ribosome display selections against the peptide neurotensin. Highly specific peptide binders were enriched after four rounds of selections using ribosome display. Four peptide side chains were shown to contribute most of the interaction energy, and single alanine mutants could be discriminated. Thus, designed ArmRP libraries can become valuable sources for peptide binding molecules because of their favorable biophysical properties and with a potential for application in general modular peptide recognition.

Introduction

In the past two decades, numerous protein scaffolds[1–3] were explored for the generation of designed binding molecules using both rational and combinatorial approaches. However, no generic peptide-binding scaffold with high specificity and affinity has yet been reported. Most importantly, no attempt has been made to exploit the modular structure of extended peptides, which contrasts with the idiosyncratic surface of a folded protein.

The recognition of extended regions of proteins, unfolded proteins or peptides from a protein digest would be particularly useful for applications in proteomics, enabling, for example, protein or mutant identification by protein chip technology or affinity chromatography.[4,5] Moreover, many posttranslational modifications (e.g., phosphorylation, acetylation and methylation) are within extended regions of proteins.

The binding of a flexible peptide chain is accompanied by a loss in entropy. This energetic cost must be compensated by the formation of specific interactions. The ability to form hydrogen bonds or other specific interactions (salt bridges or hydrophobic contacts) between the protein and the peptide side chains is therefore crucial to achieve significant binding affinities and selectivity.

A number of protein-binding scaffolds exist in nature. Antibodies are widely used to bind peptides

and have been well characterized.[6–8] Nonetheless, anti-peptide antibodies made by classical immunization or by recombinant technologies are not always of high specificity and affinity. Although peptide-binding antibodies have certain structural features in common, the orientation of the peptide is not conserved. Thus, information gained from structures of antibody–peptide complexes cannot be easily extended to generate new peptide-binding antibodies or peptide-binding proteins.

Natural peptide-binding protein domains (e.g., SH2,[9] SH3,[10,11] WW[12] and PDZ[13–15]) have been mutagenized to derive species with novel binding specificities. The peptide binding in these small adaptor domains usually depends on a particular sequence feature, for example, a phosphorylated amino acid, a stretch of proline residues or a free C-terminus. These domains thus represent solutions to specific recognition problems but seem to lack the possibility to be general peptide-binding scaffolds. Since they recognize only short amino acid stretches, generally characterized by a low affinity, it would be difficult to extend the sequence specificity. While several such domains could be fused together by flexible linkers generating some avidity, the entropy loss upon binding of such flexibly linked constructs

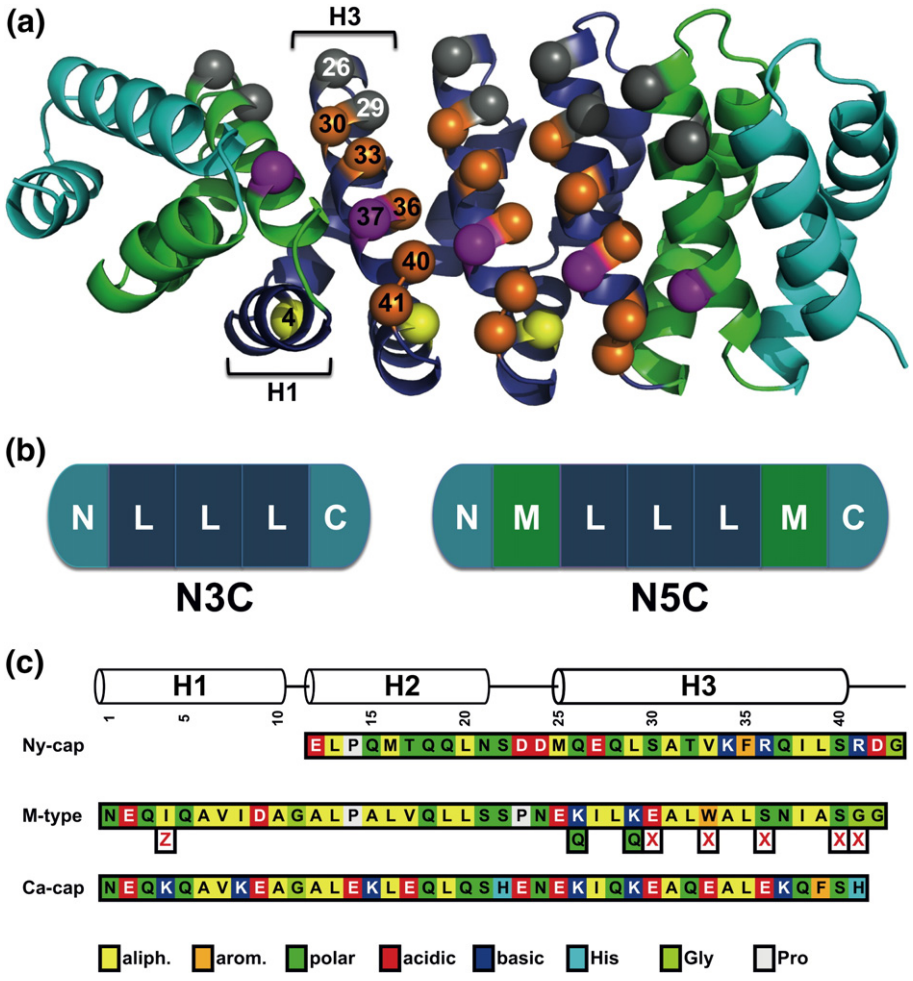

Fig. 1. Models of designed ArmRP libraries. (a) The model of an N5C molecule contains N- and C-capping repeats (in cyan), two consensus repeats (in green) and three central library modules (in blue). Residues which are randomized using all amino acids, except cysteine, glycine and proline, are indicated by orange spheres and labeled with the position within the repeat. Position 4 has a restricted randomization (EHIKQRT), and it is shown as a yellow sphere. Position 37, depicted in magenta, hosts a conserved asparagine, which is responsible for binding the target peptide backbone in available structures. Gray spheres indicate positions 26 and 29, where lysine and glutamine were allowed (variants KK, KQ, QK and QQ). The model was based on the crystal structure of a consensus-designed ArmRP[28] (Protein Data Bank ID 4DB6) and realized with PyMOL (Schrödinger, LLC). (b) Schematic representation of modules present in N3C and N5C libraries: randomized library modules (L, in blue), terminal capping repeats (N and C in cyan) and constant consensus M-type modules (M, in green). (c) Sequences of the capping repeats and the internal module M are indicated with the randomized position labeled by the red X underneath, symbolizing all 20 amino acids except Cys, Pro and Gly. The red Z in position 4 symbolizes a randomization only to Glu, His, Lys, Arg, Ile, Gln or Thr. The helix number is indicated.

would not necessarily lead to the high affinity that a larger rigid binder can provide.

The major histocompatibility complexes (MHC I and MHC II)[16] possess sufficient intrinsic variability and the ability to recognize a broad range of peptides, but difficulties in their functional expression and slow equilibration with peptide reduce their attractiveness as a scaffold candidate.

Natural repeat proteins have evolved to mediate a wide range of protein–protein interactions across all cellular compartments and across all kingdoms.[17,18] Designed repeat proteins have been developed over the past few years to facilitate the development of specific binding proteins and thus greatly expand the range of applications beyond what is possible with classical monoclonal antibodies and recombinant antibody fragments (reviewed in Ref. 2).

Several repeat proteins bind peptides, such as HEAT repeats,[19] Armadillo repeats (ArmRs)[20–23] or TPR repeats.[24] We found Armadillo repeat proteins (ArmRPs) of particular interest, since a series of Asn side chains is involved in binding and thus constraining the peptide's extended main chain through extensive hydrogen bond formation.[25–27] Every ArmR is composed of three α-helices, named H1, H2 and H3, arranged similarly in a spiral staircase, and several repeats stack to form the compact domain (Fig. 1a). Specialized repeats are present at the N- and C-termini of the protein, protecting the hydrophobic core from solvent exposure. ArmRPs consist of different families, including importins and β-catenins,[25–27] and the binding seems to be conserved along the surface generated by adjacent helices 3, almost perpendicular to the axis of helix 3. The peptide is bound to the ArmRP in an antiparallel arrangement, and protein and peptide together form a double helix that, however, unlike the DNA double helix, is very asymmetric.

The modular structure of ArmRPs might provide the basis for modular peptide recognition, in a first approximation with a dipeptide per repeat, taking advantage of the regularity of the binding site. The modularity and the conserved binding mode were the main reasons behind the choice of this protein family as the starting framework.

Nonetheless, the generation of suitable designed ArmR modules was a prerequisite in order to exploit the modular arrangement. An initial consensus-based design led to well-expressed and stable proteins that unfortunately were either dimeric or molten globules. A stable, well-expressed monomeric protein, consisting of identical internal repeats, was finally obtained using a molecular-dynamics-based approach for the stabilization of the hydrophobic core of the molten globule variant.[29]

Having the scaffold in order to obtain specific binders, we need to design and generate suitable libraries that are randomized such that they maximize the diversity of the binding surface and, at the same time, keep the stable core intact and minimize repulsive interactions within the protein surface. We then need to apply in vitro selections to test the performance of the system.

The work described here represents only the first step of a long-term project, the generation of sets of specific binders for modular peptide recognition. Here, we describe the construction, analysis and selections from a first set of such libraries by using ribosome display and the characterization of a binder for residue-specific recognition and affinity.

Results and Discussion

General considerations for generation of a designed Armadillo library

Positions responsible for peptide binding in every repeat of ArmRPs have been identified from the crystal structures of complexes of ArmRPs with their targets[21,25,29] (Supplementary Data Fig. S1) and mapped onto a consensus-designed ArmR named M-type (Fig. 1). All the positions involved, except position 41, are located on helices. Position 4 is located on helix H1 and the others are on helix H3. Recognition of the target by residues situated in secondary structure elements is characteristic of solenoid-like repeat proteins, in contrast to antibodies and some other alternative scaffolds,[1–3] where the binding site is mainly formed by loops. In the Armadillo proteins, the rigidity of these elements and the regular spacing allow the formation of the desired general binding mode that we would like to obtain in our library.

In some of the best-characterized Armadillo–peptide complexes, and especially in the minor nuclear localization sequence (NLS) binding site of importin α,[21,25] Asn in position 37 is responsible for the recognition of the target backbone by two hydrogen bonds to the peptide group of the bound oligopeptide; this asparagine is already present in the consensus and will therefore be kept constant. Positions 4, 30, 33, 36, 40 and 41 of the M-type consensus module were chosen for randomization. Position 4 contributes both to the target binding and to the hydrophobic core formation; the residues at this position will be limited to the few amino acid types compatible with the core packing. The residues appearing in natural proteins have been included, with the exception of Pro, which could destabilize the helix H1, and Ala, which is too small to form contacts with the side chain of the target peptides. Glu has been added for the potential ability to interact with positive charges, even if not present in the list of residues in natural Armadillo proteins. Positions 26, 29 and 30 are used

alternatively for the recognition of long target side chains: position 29 in the catenin subfamily and position 30 and once position 26 (repeat 5 of yeast karyopherin α) in the importin subfamily.[30] Among them, we have only selected position 30 for randomization, being most frequently used in forming side-chain contacts and considering that the importin binding is generally more regular than the catenin binding mode. Position 33 contains a Trp in the consensus and in the importin subfamily. The space occupied by the side chain allows almost any residue at this position. Position 36 shows high variability, even though small residues are preferred in importins. Residues at this position could contribute significantly to binding. However, several residues could disrupt the backbone binding of the neighboring Asn37, as observed in some catenin complexes. Position 40 displays high variability, but large side chains are often present. Position 41 is part of the loop connecting H3 with H1 of the following repeat. It interacts with the side chains of the target peptide mainly by backbone hydrogen bonds. However, several residues could be accommodated at this position providing new types of interactions.

Designed ArmRP library generation

Designed ArmRP libraries were prepared by PCR assembly of oligonucleotides (Supplementary Data Table ST1) followed by ligation. The framework was based on the M-type consensus modules.[29] Positions 4, 30, 33, 36, 40 and 41 of the M-type consensus module were chosen for randomization (Fig. 1a and Supplementary Data Fig. S2). For position 4, only Glu, His, Lys, Arg, Ile, Gln and Thr were allowed and this was encoded with a combination of three degenerate oligonucleotides. For the other randomized positions (30, 33, 36, 40 and 41), all codons except Gly, Pro and Cys were allowed and encoded by trinucleotide phosphoramidites.[31] Gly and Pro were excluded since the randomized positions are within helix H3, and Cys was excluded to avoid complications due to spurious disulfide formation. Because of the location of the randomized positions, the module sequence was shifted by six bases, compared to our original M-type assembly (Supplementary Data Fig. S2).[29]

The theoretical diversity of one designed ArmR module is 9.9×10^6, having six randomized positions. Ligation of several modules would thus potentiate the theoretical diversity accordingly, but the actual diversity will be restricted to the number of molecules present in the library. The designed ArmR library modules were PCR-assembled and subcloned for sequence analysis; 17 out of 23 single ArmRP modules analyzed (74%) showed no error and were in-frame (Supplementary Data Fig. S3). In total, 2.2×10^7 clones for a single library module were obtained, a 20-fold oversampling of the theoretical repeat module diversity.

Single library modules, which had been stored as plasmids in transformed *Escherichia coli*, were prepared on large scale, PCR-amplified and used for assembly without additional frameshift selection. The stepwise assembly of the library encoding whole ArmRP protein domains, by using type IIS restriction enzymes and amounts of DNA to always oversample the module diversity, is shown in Supplementary Data Fig. S4; from the estimated number of molecules after each ligation step, the practical library size with three randomized repeats (N3C library) was approximately 10^{11}.

The library modules were assembled as described previously[29] and then cloned into the vector pPANK.[32] In subsequent versions of the library (see below), the vector already contained the "constant" N- and C-terminal capping repeats in order to maximize the quality of the library. A final "N3C library" was obtained (Fig. 1b), the nomenclature reflecting three internal randomized designed ArmR modules between the N-terminal capping repeat and the C-terminal capping repeat. By sequencing of randomly picked clones, we found 13 out of 20 (65%) N3C ArmRPs to be in-frame and having no stop codon, while with the N5C library (see below), 14 of 34 (41%) were in-frame and had no stop codon.

From the amount of DNA and the estimated number of molecules after each ligation step, corrected by the fraction of in-frame clones, the library size can be estimated as greater than 10^{11}. This diversity can be covered in a typical ribosome display selection,[33] and it will even be further increased by random mutations due to PCR errors.

Further backbone versions of the library

The original M-type design contains Lys at positions 26 and 29, but earlier experiments[34] had uncovered an electrostatic repulsion. Thus, both positions were converted, individually or together, from Lys to Gln. The original design is thus termed the KK-type, while the new variants are termed KQ, QK and QQ (Fig. 1a). The QK version led to aggregating molecules and was not pursued further. Biophysical studies revealed that the QQ mutants displayed the most favorable properties.[34] The Lys→Gln mutations at positions 26 and 29 were introduced during new library module assembly using primers carrying the mutations, and the product was amplified with outer primers as above.

Stabilizing the library domain by flanking constant consensus modules (N5C library)

The analysis of library members of the originally assembled N3C libraries showed that rather few molecules behaved as monomers in size-exclusion

chromatography (SEC) and had low 8-anilino-napthalene-1-sulfonate (ANS) affinity, which would have indicated a stable structure as opposed to a molten globule (Fig. 2a and c). This is in sharp contrast to the original consensus sequence, which had none of these problems.[34] Through the introduction of randomized residues, factors such as lower helical propensity, electrostatic repulsion and exposed hydrophobic residues can all contribute to poorer biophysical properties, compared to the full consensus molecules.

To stabilize the whole library, we ligated consensus modules (M) flanking the randomized modules to result in Ny-M-L-L-L-M-Ca molecules (N5C for short), where Ny is the yeast-derived N-capping repeat, M is the non-randomized "M-type" consensus module, L is the randomized library module, and Ca is the artificial C-terminal capping repeat[29] (Fig. 1b). This approach was chosen since repeat protein stability increases with length, as observed for Armadillo, Ankyrin and Tetratricopeptide repeat proteins.[29,35–37]

The construction of this N5C library is described in Supplementary Methods and shown in Supplementary Data Fig. S5; the amount of DNA was chosen high enough to be sufficient to always oversample the diversity of the designed ArmR module library.

Evaluating the optimal ArmRP format for library selections

We thus had to consider two parameters to determine the optimal format of the library, the size of the protein (N3C versus N5C) and the backbone charges in positions 26 and 29 (KK, KQ and QQ, with QK already found unsuitable). Data are summarized in Supplementary Data Table ST2. SEC indicated

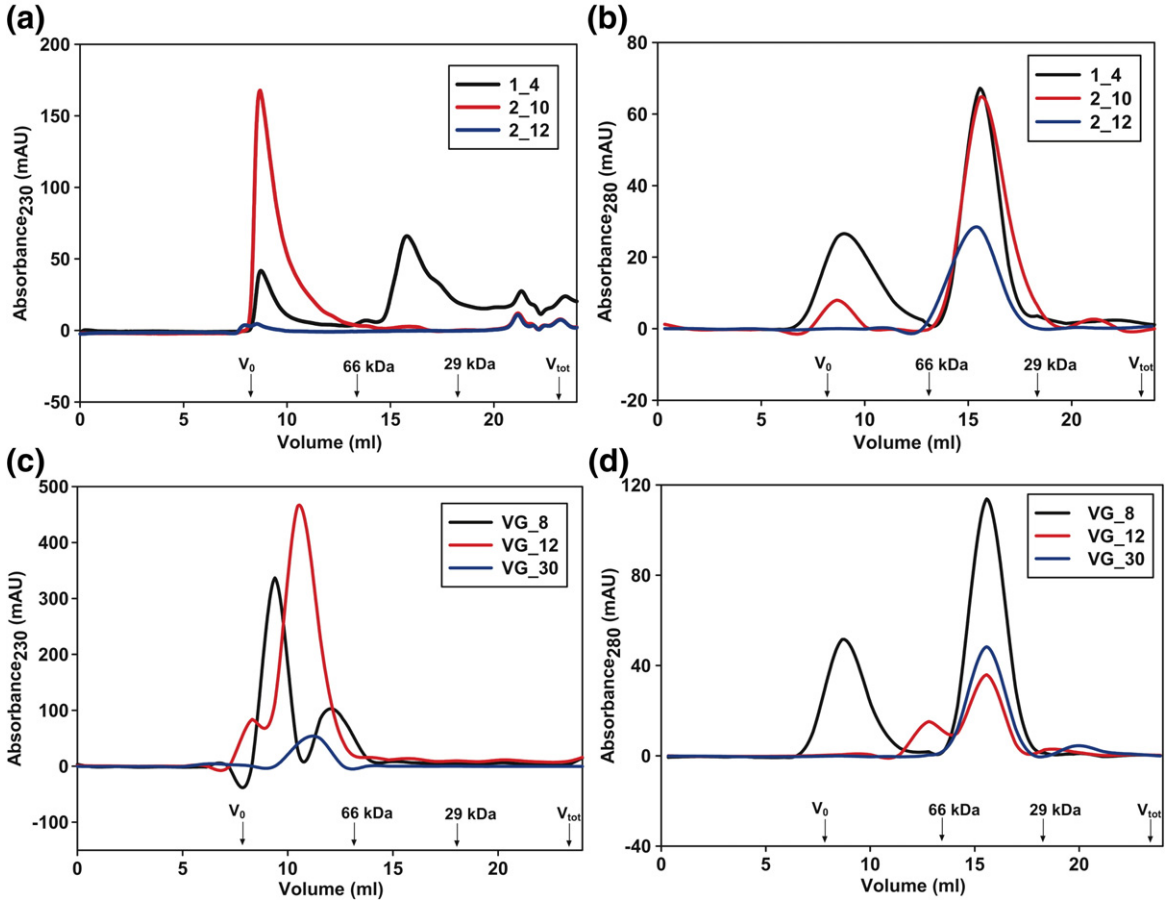

Fig. 2. SEC of library members of the N3C KK library: (a) unselected members and (c) selected members form an early test selection. They do not show any monomeric peak but elute as mixtures of oligomers and aggregates. (b and d) SEC of the same library members as shown in (a) and (c), respectively, after they had been converted to the N5C format, having constant M-type module flanking the randomized library modules (cf. Fig. 1). All the proteins in N5C format show at least a monomer peak. V_0 indicates the void volume, and V_{tot} indicates the total volume of the column. BSA (MM = 66 kDa) and carbonic anhydrase (MM = 29 kDa) were used as molecular mass markers, and the corresponding elution volumes are indicated by the arrows. The experiment was performed in 20 mM Tris and 50 mM NaCl (pH 8) with a Superdex 200 column. The elution was followed by absorbance at 230 nm for N3C library members, since some of the unselected members lack tryptophan and do not show absorption at 280 nm.

that 9 out of 11 KK-type library members (6 converted from N3C clones and 5 randomly picked N5C clones) display aggregation tendencies to variable degrees (Figs. 2b and d and 3a). KQ library members eluted entirely as monomers or as a mixture of monomer and oligomers, but they still formed soluble

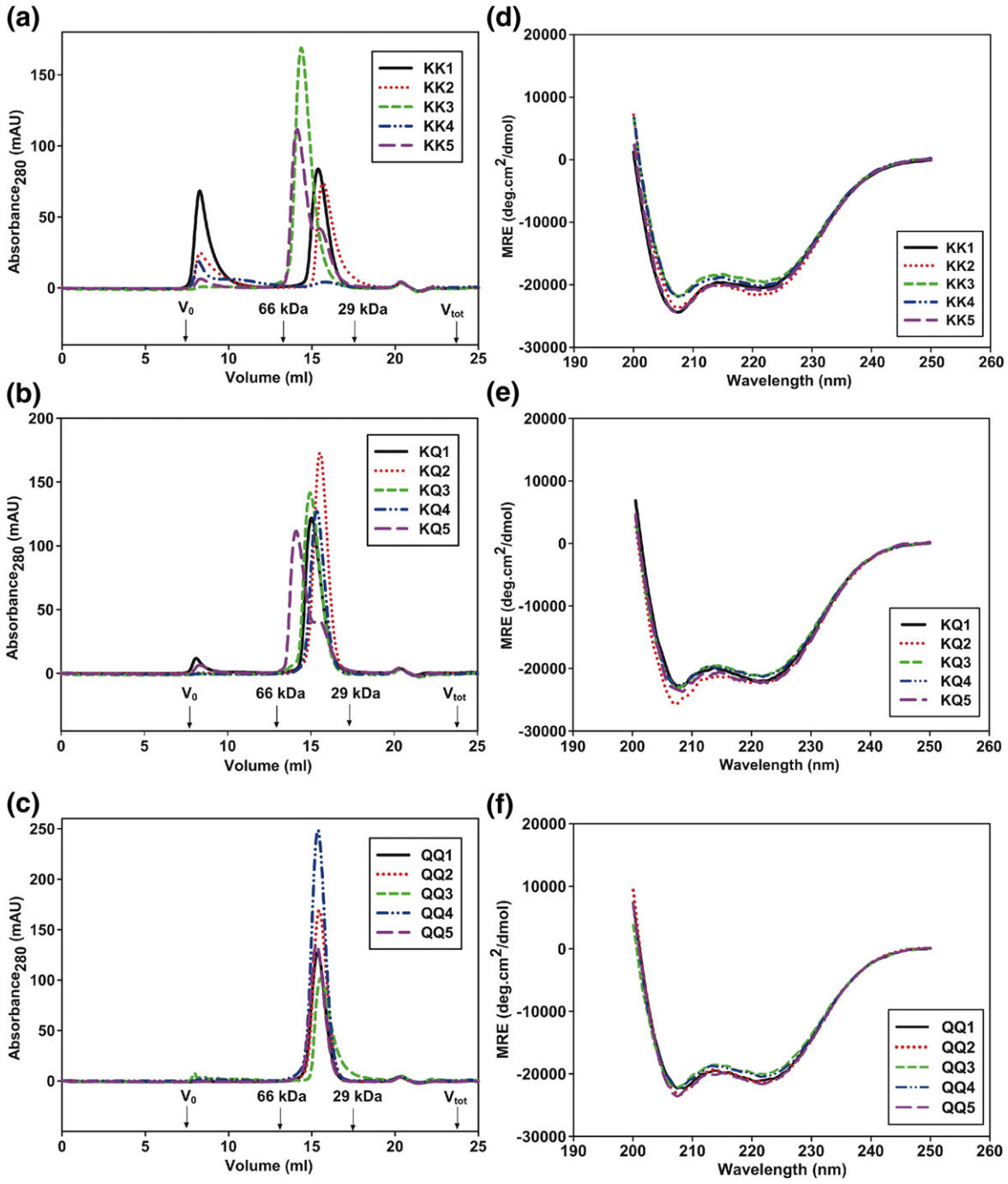

Fig. 3. Biophysical characterization of unselected N5C library members from designed ArmRPs. The N5C libraries differ in the residue at positions 26 and 29 and are named accordingly as KK, KQ and QQ (the QK library was not pursued further) (see the text). SEC of KK (a), KQ (b) and QQ (c) library members containing three randomized internal modules flanked by constant consensus modules of type M and terminal N- and C-capping repeats (Fig. 1). V_0 indicates the void volume, and V_{tot} indicates the total volume of the column. BSA (MM = 66 kDa) and carbonic anhydrase (MM = 29 kDa) were used as molecular mass markers; the corresponding elution volumes are indicated by the arrows. The experiment was performed in 20 mM Tris and 50 mM NaCl (pH 8) with a Superdex 200 column. CD spectra of the corresponding KK (d), KQ (e) and QQ (f) library members. The values are reported as MRE.

aggregates in minor proportions (Fig. 3b). The broad peaks observed for some KQ library members is due to a mixture of dimers and monomers, as indicated by multi-angle light scattering (MALS) (data not shown). For those proteins forming a mixture of aggregates and monomers, the monomeric peak after separation remains monomeric, as shown by reanalyzing it by gel chromatography.

In contrast, the QQ-type library members were eluted as monomers or predominantly monomeric (Fig. 3c), with approximately the same elution volume of full consensus YM_5A (consisting of five internal M-type modules) (see below). Some of these N5C library members still show a small peak at earlier and void volumes, indicating the presence of a small proportion of oligomers and soluble aggregates.

It is remarkable that N3C library members of the KK- and KQ-types, which showed oligomeric peaks in SEC (shown for the KK clones in Fig. 2a and c), could be shifted to either complete monomers (VG_30) or at least a mixture of monomers (major fraction) and aggregates (minor fraction), when converted to the N5C format (Fig. 2b and d) by the addition of flanking consensus M-type modules. This clearly demonstrates the benefit of adding these flanking modules.

The monomeric N5C members elute earlier than expected for a globular protein with the same mass (about 31.5 kDa); however, MALS carried out with several members showed them to be monomeric (data not shown), and this elution behavior is consistent with the elongated shape of the Armadillo scaffold.

Two further characterizations did not discriminate between members of different libraries: CD spectra, recorded for all library members using their monomeric fractions, were found to be remarkably similar among all the library members and for YM_5A, indicating that no significant differences were detectable at the level of secondary structure.

Furthermore, all of the above library members that were found to be correct at the DNA level could be expressed in soluble form in large amounts in *E. coli* (about 80 mg/l culture), similar to YM_5A (Supplementary Data Fig. S6), and ran at the expected MM (*m*olecular *m*ass) position in 15% SDS-PAGE.

In summary, these experiments revealed that the best library format was the N5C format in conjunction with the QQ subtype, as this has the highest proportion of soluble, monomeric members, showing single symmetric peaks in SEC, identical with full consensus YM_5A. These experiments also demonstrated the feasibility of obtaining additional stabilization by adding flanking repeats to designed ArmRP libraries.

Choice of the target

To test the library for its binding properties, the chosen target was a peptide that contained amino acids of different types, had no structure but was long enough to be able to bind in different registers. Neurotensin (NT), a 13-mer peptide with the sequence QLYENKPRRPYIL, was selected as target, as its random-coil nature was shown by solution NMR.[38] Nonetheless, the presence of prolines might lead to a binding mode different from that of NLS to importin α.

To focus the selection on the peptide and away from any carrier or capturing protein (such as streptavidin), we prepared the target in two different forms. First, NT as a synthetic peptide was linked at its N-terminus via an amide bond to 6-aminocaproic acid, extended by two β-alanines and biotin (Supplementary Data Fig. S7a). Biotinylated NT thus maintains a free carboxyl group at the C-terminus. Second, the peptide was expressed as fusion protein [fusion to the C-terminus of phage λ protein D (pD)]. The fusion protein contains at its N-terminal end a histidine tag for purification and a biotinylated Avi tag for immobilization. A linker region keeps the target peptide distant from the core of the fusion partner, preventing steric hindrance (Supplementary Data Fig. S7b). To focus selection on the last eight amino acids of NT, we prepared an additional pD fusion construct for prepanning: the construct was identical, except that the C-terminal NT sequence was modified to QLYENAAAAVVVV (named pD_NT_A_4V_4 for short).

Strategies for selections

Ribosome display was chosen as selection method for its capability of handling large libraries, with up to 10^{12} members. Three different strategies (Table 1) were adopted during selections in order to select highly specific peptide binders that recognize and

Table 1. Scheme of the RD selections using an N5C (QQ) designed ArmRP library

Round no.	Immobilizing surface	Target concentration (NT peptide) (μM)	SI[a]	SII[b]	SIII[c]	Washing	No. of RT PCR cycles
1	Neutravidin	1	NT	NT	pD_NT	6× with 1-min intervals	40
2	Streptavidin	0.25	NT	pD_NT	pD_NT	5× with 2-min intervals	35
3	Neutravidin	0.1	NT	NT	pD_NT	5× with 4-min intervals	30
4	Streptavidin	0.075	NT	pD_NT	pD_NT	5× with 5-min intervals	25

[a] Strategy I: only synthetic peptide was used as target in all selection rounds.
[b] Strategy II: synthetic peptide was alternated with peptide attached to fusion pD during selections.
[c] Strategy III: only peptide attached to fusion pD was used.

Table 2. List of target peptides used in ELISAs

Name	Biotinylation	Linker	Fusion partner	Sequence
NLS	Yes	Ttds[a]	—	KKKRKV
Cro	Yes	Ttds	—	PRTSSF
pDCro	Yes	Ttds	pD	PRTSpSF[b]
NT	Yes	LC[c] and $(\beta\text{-Ala})_2$	—	QLYENKPRRPYIL
pD_NT	Yes	GS[d]	pD	QLYENKPRRPYIL
pD_NT_A_4V_4	Yes	GS	pD	QLYENAAAAVVVV
ERK	Yes	Ttds and $(\beta\text{-Ala})_5$	—	DHTGFLTEYVA
pERK	Yes	Ttds and $(GGS)_2$	—	DHTGFL-pT-E-pY-VA[b]
pD_NT_(8–13)	Yes	GS	pD	RRPYIL
NT_(8–13)	No	—	—	RRPYIL
NT_free	No	GS	—	QLYENKPRRPYIL

[a] Ttds, 4,7,10-trioxa-1,13-tridecanediamine succinimic acid.
[b] pS, phosphoserine; pT, phosphothreonine.
[c] LC, 6-aminohexanoic acid; connected to biotin.
[d] Gly-Ser.

bind only the peptidic sequence of the NT peptide. In all cases, selection against a biotinylated target was used. The target was either a synthetic peptide or a fusion protein with pD (see above) (Supplementary Data Fig. S7a and b), and it was either coupled to microtiter wells via neutravidin or bound to streptavidin linked to magnetic beads. The procedures differed in how targets and immobilization were alternated. Selections were based on the standard protocol for ribosome display[39–41] as adapted for the DARPin library.[42]

In the first strategy (SI), the synthetic peptide was used in all selection rounds. Prepanning was performed first against a neutravidin-coated surface, followed by prepanning on the unrelated ERK peptide (Table 2) bound to neutravidin, which contains a similar linker region but differs only in the peptide sequence.

In the second strategy (SII), the library pool was alternatingly exposed to the synthetic peptide or peptide fused to pD in successive selection rounds. When using peptide, we performed prepanning against a neutravidin-coated surface, followed by prepanning on ERK peptide bound to neutravidin. When the pD fusion protein (pD_NT) was used, a prepanning was performed against streptavidin beads, followed by a pD fusion peptide differing in the last eight amino acids (pD_NT_A_4V_4).

In the third strategy (SIII), the library pool was always exposed to peptide fused to pD in all rounds. Prepanning was performed against neutravidin or streptavidin, followed by prepanning against pD_NT_A_4V_4.

The selection cycles were performed with increased washing stringency (Table 1), and the result of the RT (reverse transcription) PCR showed progressive enrichment across the selection rounds (Supplementary Data Fig. S8).

Screening of single clones

In the second strategy (SII), where the form of the target (peptide or pD fusion) was changed in every round, the highest enrichment was observed. The selected pool was cloned in the vector pPANK and single clones were assayed for NT binding in a 96-well format (Supplementary Data Fig. S9). Clones, which showed strong binding signal to NT peptide in ELISA, gave rise to proteins running at the expected size in SDS-PAGE (Supplementary Data Fig. S10).

Out of 30 clones sequenced, 29 clones had an identical sequence, and the remaining 1 clone differed in a single amino acid at position 116. These two binders were termed as VG_328 and VG_306 (Fig. 4). Both proteins were purified by immobilized metal-ion affinity chromatography (IMAC), and recognition of the peptide (but not streptavidin or neutravidin) was verified by ELISA (Fig. 5), with the binding of VG_328 giving twice the signal of VG_306. The full consensus protein YM_5A did not interact with NT, indicating that the consensus-designed (non-randomized) ArmRP scaffold per se does not bind NT.

The sole difference between VG_306 and VG_328 is a single point mutation, Tyr116 to His, located in the first library repeat L1 of this Ny-M-L1-L2-L3-M-Ca molecule, a randomized position on helix H3 at position 30. As this mutation, located in the middle of the randomized surface, influences binding strength, the peptide can be assumed to bind to the expected surface containing the randomized residues, even though the exact location must await the structure determination of the complex.

Biophysical characterization of binder VG_328

The two selected binders VG_306 and VG_328 were expressed and purified by IMAC, and the mass was verified by mass spectrometry. The yield was identical with that of consensus-designed ArmRP YM_5A (Supplementary Data Fig. S11).

Binder VG_328 displays high ANS binding, compared to consensus-designed ArmRP YM_5A (Fig. 6a). We do not know whether this is due to the presence of hydrophobic residues on the

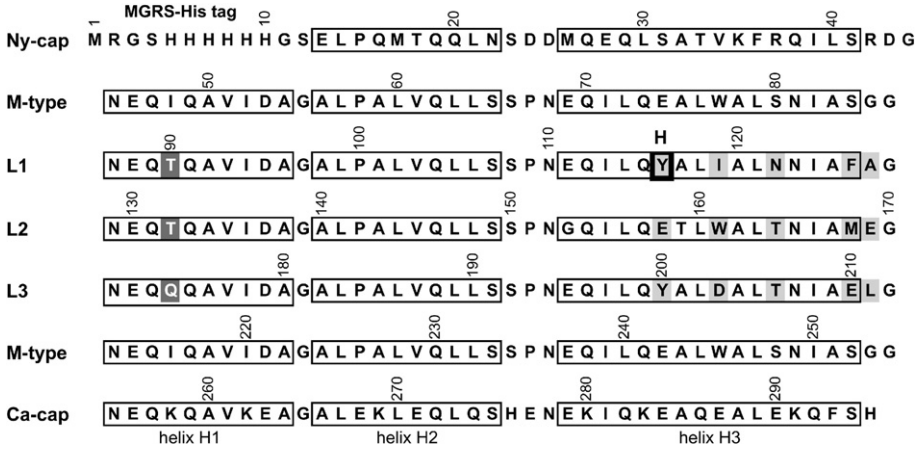

Fig. 4. Sequence of the selected binder VG_328, obtained from the N5C library (cf. Fig. 1b). The sequence is consecutively numbered. The library modules are indicated as L1–L3. The randomized positions are shown with black letters on light-gray background (all 20 amino acids except Cys, Pro and Gly) or white letters on dark-gray background (Glu, His, Lys, Arg, Ile, Gln or Thr). At position 116 (dark frame), the Tyr→His change is shown, the sole difference to VG_306, which binds with lower affinity (Fig. 5).

randomized surface or transient access to internal pockets in the structure. Other parameters are inconsistent with a molten globule state and indicate a native-like state, including NMR studies that demonstrated high-quality ^{15}N, ^{1}H correlation spectra (C. Ewald, R. Watson and O. Zerbe, unpublished results).

Gel filtration, mass spectroscopy and MALS (data not shown) detected only monomers and could not confirm the presence of any other protein species (Fig. 6b). For both VG_328 and YM_5A, the MM estimated from gel filtration is slightly higher (by a factor of 1.25) than the value calculated from globular standard proteins (Table 3) (cf. Figs. 2 and 3) (see above), consistent with the elongated shape of the ArmRPs in combination with a flexible N-terminal tail (MRGSHHHHHHGS).

The CD spectra of the IMAC-purified VG_328 and consensus-designed YM_5A can almost be superimposed (Fig. 6c). The CD signal at 222 nm was chosen to monitor stability during thermal and denaturant-induced unfolding. VG_328 shows a cooperative transition in both unfolding studies. The midpoint of transition during thermal denaturation (T_m) of VG_328 is 74 °C, which is still high but somewhat lower than the 80 °C measured for consensus-designed YM_5A (Fig. 6d). We also investigated unfolding induced by guanidinium chloride (GdmCl) (Fig. 6e). For VG_328, a midpoint of transition of 3.3 M GdmCl is observed, which is somewhat lower than that of YM_5A, with approximately 4.2 M.

The lower stability of VG_328 compared to YM_5A might be due to randomization, where one of the randomized positions (position 4) of each library repeat module contributes potentially both to peptide binding and to hydrophobic core packing. Also, a point mutation (E153G) was acquired at the beginning of helix H3 of the third internal repeat of VG_328 (second randomized library repeat module). Nonetheless, combining the CD data with the findings from the protein expression and gel-filtration experiments, it can be concluded that the selected binder VG_328 is a soluble, monomeric protein, consistent with having an Armadillo fold as designed.

Specificity of binding

The selected binder VG_328 was purified and its specificity was evaluated using ELISAs, by testing

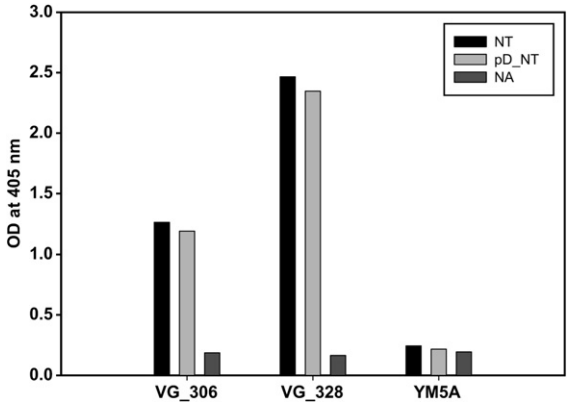

Fig. 5. Target recognition by selected ArmRPs VG_306 and VG_328. ELISA analysis of purified proteins for binding to peptide targets in different forms. NT, pD_NT and NA indicate the target peptides, NT, and NT fused to pD bound to neutravidin via biotin and neutravidin alone, respectively. YM_5A is included as a nonbinding control. The binding of the designed ArmRPs was detected using a primary anti-RGSHis antibody and an anti-mouse antibody coupled to alkaline phosphatase as secondary antibody (Qiagen). The signal was developed for 2 h after addition of the substrate.

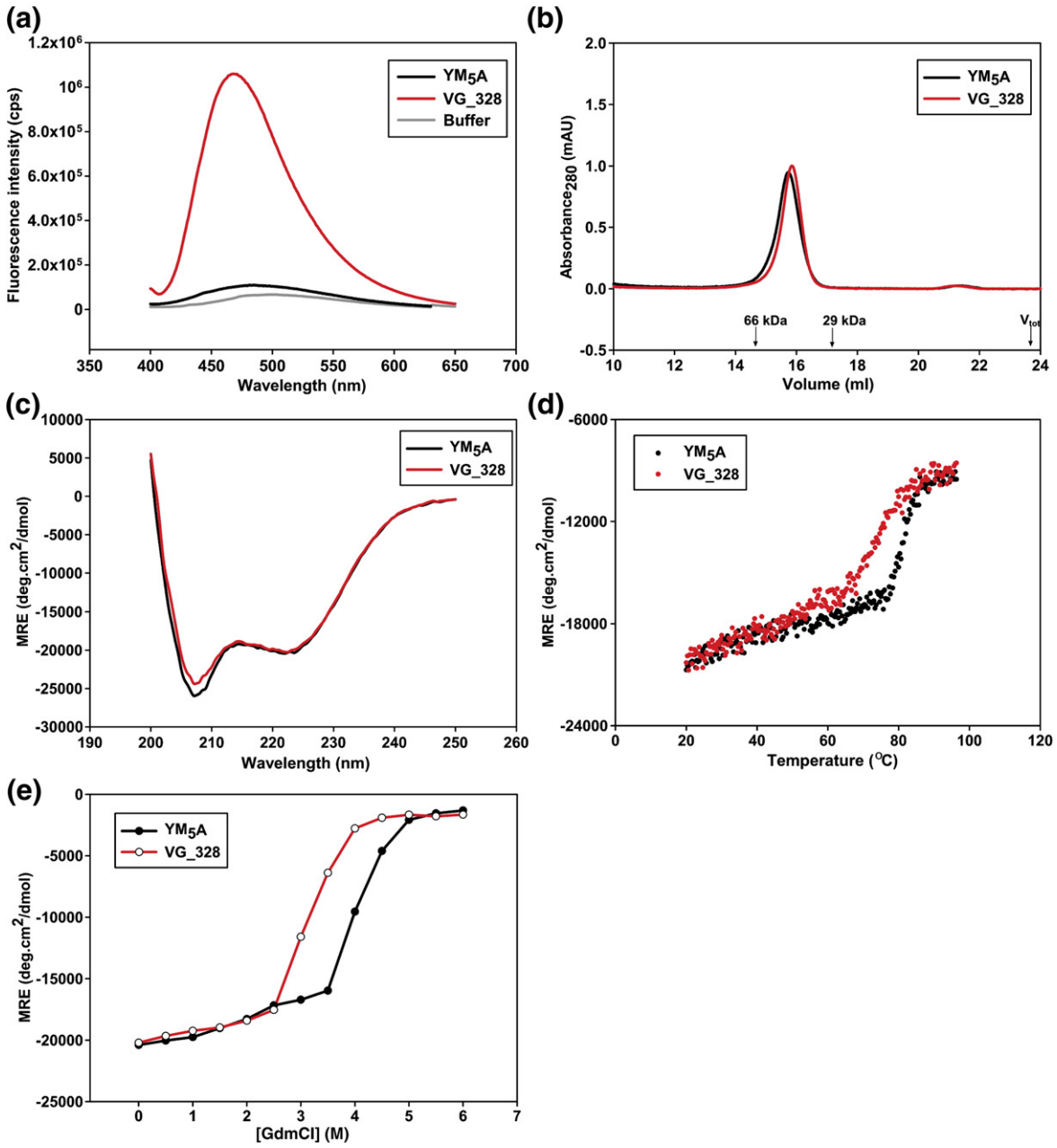

Fig. 6. Biophysical characterization of the selected binder VG_328. Consensus-designed protein YM_5A is included as reference. (a) ANS binding of VG_328 shows higher fluorescence than full consensus protein YM_5A. The values without buffer subtractions are shown. (b) SEC of VG_328 and YM_5A. V_0 indicates the void volume, and V_{tot} indicates the total volume of the column. BSA (MM = 66 kDa) and carbonic anhydrase (MM = 29 kDa) were used as molecular mass markers; the corresponding elution volumes are indicated by the arrows. The experiment was performed in 20 mM Tris and 50 mM NaCl (pH 8) with a Superdex 200 column. (c) CD spectra, (d) thermal denaturation curves and (e) guanidinium-chloride-induced denaturation curve of VG_328. The denaturation experiments were followed by CD. The values of MRE at 222 nm are reported.

binding to 10 different peptides differing in sequence from the NT peptide (Table 2). VG_328 displayed strong binding to NT and no binding to other peptides tested over background (Fig. 7a). A very weak binding to an NLS peptide was observed, which may be due to a positively charged cluster similar in the NT target peptide.

To confirm the specificity of VG_328, we used unbiotinylated NT peptide or unbiotinylated pD_NT as competitors. Clear competition was observed,

Table 3. Biophysical properties of consensus YM_5A and binder VG_328

Constructs[a]	Residues (repeats)[b]	pI[c]	MM_{calc} (kDa)[d]	Oligomeric state[e]	MM_{obs} (kDa)[f]	$MM_{obs/calc}$[g]	CD_{222} (MRE)[h]	Observed T_m (°C)[i]	CD GdmCl (M)[j]
YM_5A	295 (7)	4.4	31.5	Monomer	38.6	1.23	−20,435	80	4.2
VG_328	295 (7)	4.3	31.2	Monomer	39.9	1.28	−20,199	74	3.3

[a] VG_328 (binder) and designed full-consensus ArmRP YM_5A.
[b] The number of residues includes the $MRGSH_6$ tag; the number of repeats includes capping repeats.
[c] Isoelectric point.
[d] Molecular mass calculated from the sequence; masses were confirmed by mass spectrometry.
[e] Oligomeric state as indicated by multi-angle static light scattering.
[f] Apparent molecular mass as determined in SEC.
[g] Ratio between observed and calculated molecular masses.
[h] MRE at 222 nm (deg cm^2/dmol).
[i] T_m observed in thermal denaturation by CD.
[j] Midpoint of transition in GdmCl-induced denaturation, measured by CD.

while none of the other peptides tested could inhibit binding (Fig. 7b). These results support the hypothesis that the selected ArmRP VG_328 binds to the NT peptide in solution and is highly specific for this sequence.

Alanine scanning of the bound peptide was used to determine the contributions of single residues to binding (Table 4). These experiments were carried out with the pD_NT fusion protein whose binding was shown to be equivalent to the synthetic peptide. All constructs were in vivo biotinylated, expressed and IMAC-purified. Binding of VG_328 only to some of the alanine mutants of pD_NT was reduced by about 50% in ELISA (Fig. 7c): positions 7 (Pro), 8 (Arg), 9 (Arg) and 11 (Tyr) are crucial for binding, whereas mutation of other residues makes no significant difference in binding.

To confirm these results, we collectively mutated all the residues that did not contribute to binding to alanine in one construct (pD_NT_15), while in another construct, we exchanged the residues that did contribute to binding to alanines all at once (pD_NT_16) (Fig. 7d). It was observed that binding of VG_328 to pD_NT_16 was completely lost, while VG_328 shows unaffected binding to pD_NT_15 (similar to pD_NT). These results confirm that Pro7, Arg8, Arg9 and Tyr11 are the amino acid side chains most critical for binding of VG_328 to NT.

As a final confirmation of binding specificity, we tested the C-terminal hexapeptide NT_(8–13). For this purpose, we generated the new peptide fusion construct pD_NT_(8–13) in which the first seven amino acids of NT were deleted. VG_328 binds to the C-terminal hexapeptide sequence of NT (Fig. 7e), demonstrating that this hexapeptide does indeed contain the key epitope, as prescribed by the panning strategy. Nonetheless, the binding strength of VG_328 was reduced compared to pD_NT. These results indicate that Pro7, which is lacking in this construct, makes a contribution to binding, consistent with the results from the alanine scan. Importantly, the free hexapeptide NT_(8–13) could serve as a competitor in solution (Fig. 7f), demonstrating that a binder against this epitope was indeed generated.

Affinity determination

To determine the affinity of the selected binder VG_328 by surface plasmon resonance (SPR) with a BIACORE instrument, we immobilized the biotinylated synthetic NT peptide on a streptavidin chip and injected increasing concentrations of the protein. The binding curves indicated rapid equilibration (Fig. 8a), and thus, the affinity could be determined from the plateau level as about 7 μM at 4 °C (Fig. 8b). While this affinity is only moderate, it should be emphasized that probably only four amino acids contribute to binding and that selection had emphasized specificity and not affinity. A wide range of affinities with single residue discrimination has also been seen in natural ArmRP, for example, in importin α.[43]

To further test the affinity to the free peptide target, albeit in an only semiquantitative approach, we carried out competition ELISA experiments with purified VG_328. Here, a constant amount of VG_328 (50 nM) was preincubated with different concentrations of free NT peptide at 4 °C for 1 h on an Eppendorf shaker. The samples were then applied to wells containing constant amount of immobilized biotinylated NT peptide (20 nM) (Fig. 8c). It was found that the binding was reduced by more than 50% when VG_328 was competed with 200 nM free NT peptide under these conditions.

Future experiments will have to systematically address strategies for routinely obtaining higher affinities yet maintaining high specificity. Since our focus is to develop a technology for modular peptide binding rather than to select individual peptide binders that might be improved through affinity maturation, the structure determination of several ArmRP–peptide complexes is pivotal to determine whether the repeat spacing of the ArmRP matches

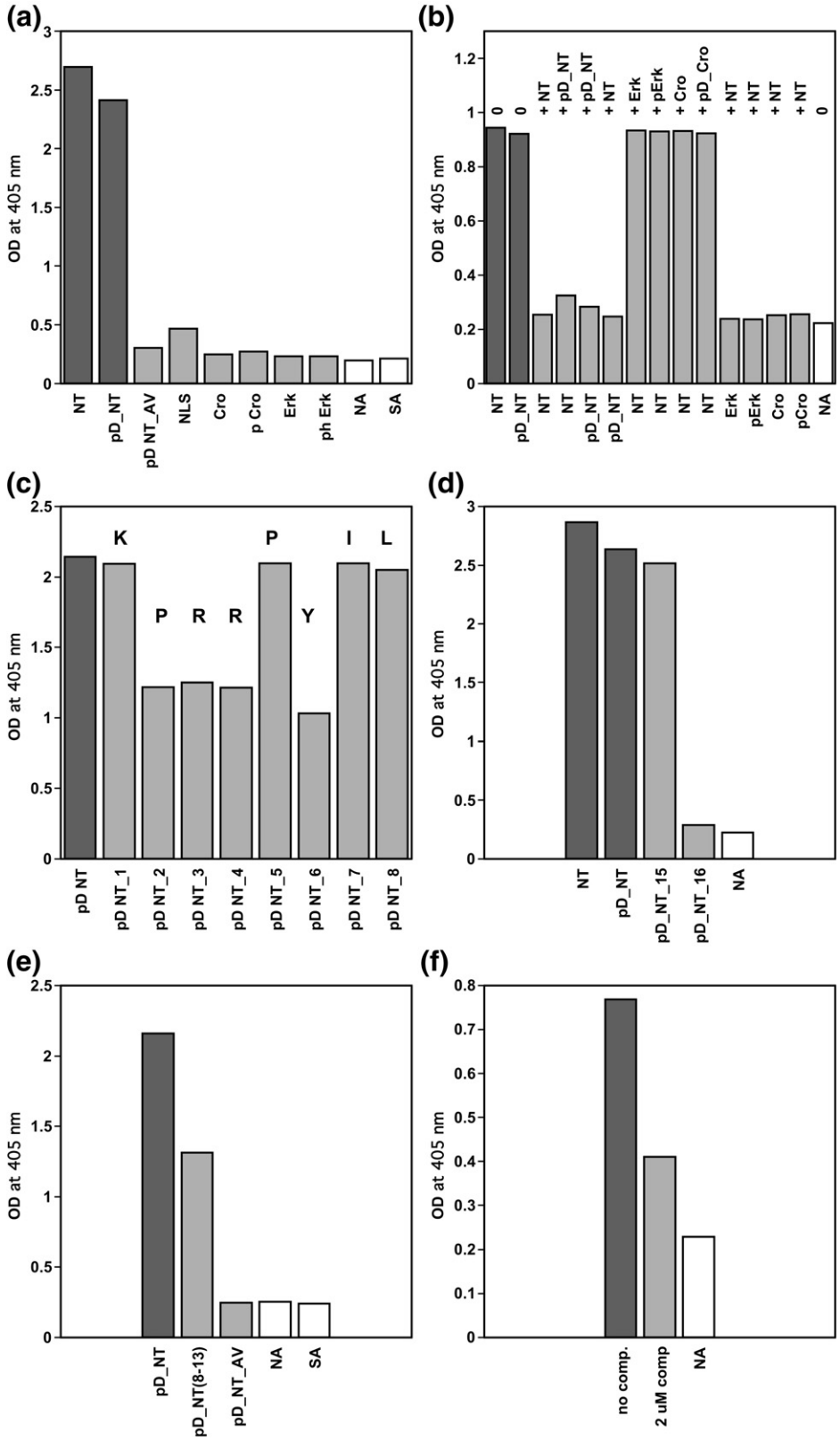

Fig. 7 (*legend on next page*)

Table 4. Single and triple alanine mutant variants of pD_NT peptide

Name	Alanine position in NT	Sequence
pD_NT	—	QLYENKPRRPYIL
pD_NT_1	6	QLYENAPRRPYIL
pD_NT_2	7	QLYENKARRPYIL
pD_NT_3	8	QLYENKPARPYIL
pD_NT_4	9	QLYENKPRAPYIL
pD_NT_5	10	QLYENKPRRAYIL
pD_NT_6	11	QLYENKPRRPAIL
pD_NT_7	12	QLYENKPRRPYAL
pD_NT_8	13	QLYENKPRRPYIA
pD_NT_9	6,7,8	QLYENAAARPYIL
pD_NT_10	7,8,9	QLYENKAAAPYIL
pD_NT_11	8,9,10	QLYENKPAAAYIL
pD_NT_12	9,10,11	QLYENKPRAAAIL
pD_NT_13	10,11,12	QLYENKPRRAAAL
pD_NT_14	11,12,13	QLYENKPRRPAAA
pD_NT_15	6,10,12,13	QLYENAPRRAYAA
pD_NT_16	7,8,9,11	QLYENKAAAPAIL

All mutants were expressed as fusions to pD and biotinylated at the Avi tag.

that of the peptide precisely enough, whether binding is in the consensus mode and whether general design features need to be improved. This knowledge will also be crucial for improving randomization strategies and extending the interaction interface. Nonetheless, well-known affinity improvement strategies to individual binders would be available already today, such as generation of avidity or the development of a sandwich binding strategy.[44]

Applications of binder VG_328

The utility of binder VG_328 was demonstrated in Western blots and pull-down experiments (Fig. 9). Crude extracts from *E. coli* expressing pD_NT and pD_NT_16 as negative controls were separated by SDS-PAGE, transferred to the membrane and detected by using VG_328 as the primary reagent. While pD_NT is clearly detected, pD_NT_16 and other *E. coli* proteins are not visible or only visible as faint bands.

For pull-down experiments, crude extracts from *E. coli* expressing VG_328 were incubated with streptavidin beads on which biotinylated pD_NT or the negative control pD_NT_16, respectively, had been immobilized. VG_328 was found to interact only with pD_NT and not with the negative control pD_NT_16 (Fig. 9b).

The additional bands visible on the gel are all expected from the experiment. Besides streptavidin (non-covalently bound subunits become dissolved after boiling in SDS) and pD_NT or pD_NT_16, bovine serum albumin (BSA) coming from the blocking steps is seen in all samples. When no biotinylated pD fusion was bound, the streptavidin beads alone were able to capture a 17-kDa protein from crude *E. coli* extracts, consistent with the biotin carboxyl carrier protein, which is the only naturally biotinylated protein in *E. coli* and has a theoretical MM of 16.7 kDa.[45,46]

In summary, binder VG_328 shows the specificity previously seen in ELISA, also in Western blot and pull-down experiments in crude *E. coli* extracts. Although its affinity is not satisfactory, VG_328 can serve to demonstrate that the specific recognition of ArmRPs can be exploited in applications such as Western blots or pull-down experiments.

Conclusions

In summary, we generated a designed ArmRP library module composed of fixed framework positions and randomized positions on the binding surface through sequence and structure consensus analyses. The most stable designed ArmRP library domains were generated by cloning three library modules flanked by a consensus module on either side and N- and C-terminal capping modules. All tested library members are well expressed in *E. coli*, can be purified easily in large amounts and are

Fig. 7. Specificity analysis of the selected binder VG_328. (a) The interaction of VG_328 (100 nM) with other immobilized peptide targets (immobilized at 50 nM) differing in peptide sequence was compared to NT. The biotinylated target peptides were bound to neutravidin. NA, neutravidin alone; SA, streptavidin alone. (b) Specificity analysis of VG_328 by inhibition ELISA. VG_328 (50 nM) was incubated with 500 nM (10-fold excess) of different free peptides (indicated above the bars) before binding on similar or other immobilized peptides (indicated below the bars). Target peptide was provided in both forms: biotinylated NT and pD_NT (immobilized at 25 nM). NA refers to neutravidin alone. (c) Alanine scanning: ELISA analysis of VG_328 for binding against various alanine mutants of pD_NT. (d) Combination of alanine mutants. pD_NT_15 refers to NT peptide fusion in which amino acids not contributing to binding have been simultaneously mutated to alanines. pD_NT_16 refers to the NT peptide fusion in which amino acids contributing to binding have been mutated to alanines (cf. Table 4). NA, neutravidin alone. (e) Binding of VG_328 to the C-terminal hexapeptide of NT (RRPYIL) fused to pD [pD_NT_(8–13)] and to a construct with the last eight amino acids of NT replaced by the sequence AAAAVVVV (pD_NT_AV) (cf. Table 2). NA and SA refer to neutravidin alone and streptavidin alone, respectively. (f) Competition with the hexapeptide RRPYIL. The fusion protein pD_NT_(8–13) was immobilized at 25 nM. VG_328 (100 nM) was competed with 2 μM free hexapeptide of NT. The binding of VG_328 was detected using a primary anti-RGSHis antibody and anti-mouse AP-coupled secondary antibody (Qiagen). The signal was developed for 2 h after addition of the substrate.

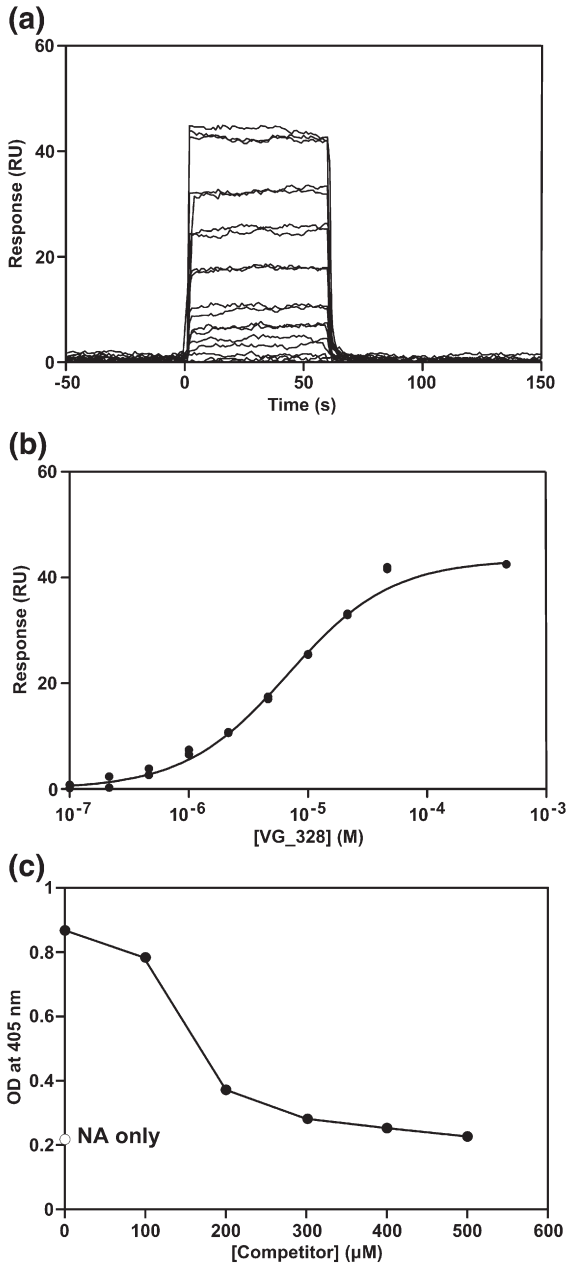

Fig. 8. (a) SPR sensograms of the selected binder VG_328 on biotinylated NT peptide, immobilized on a streptavidin chip, at 4 °C. Double referencing was applied; for details, see the text. (b) Plateau heights from (a) are plotted as a function of VG_328 concentration. A K_D of 7 µM is obtained from the fit indicated. For details, see the text. (c) Competition ELISA of the selected binder VG_328. VG_328 (50 nM) was incubated with different concentrations of free NT peptide as competitor before the mixture was added to immobilized NT peptide. The binding of VG_328 was detected using a primary anti-RGSHis antibody and anti-mouse AP coupled secondary antibody (Qiagen). The signal was developed for 2 h after addition of the substrate.

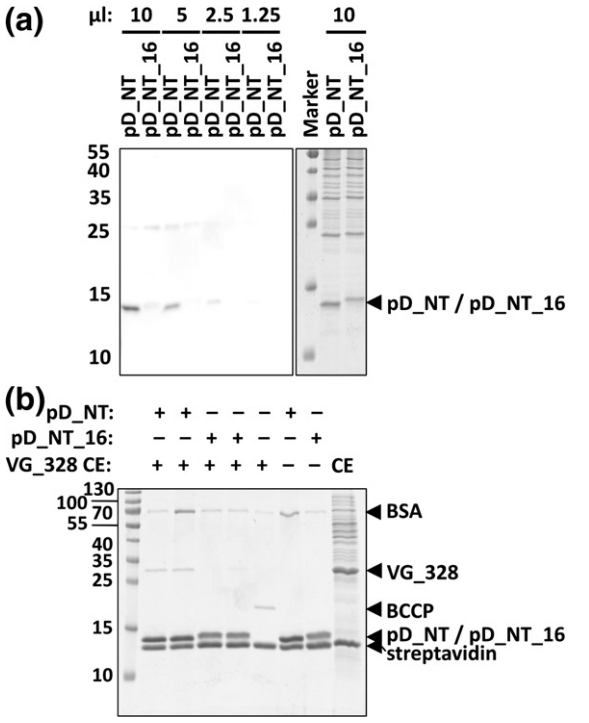

Fig. 9. (a) Western blot (left) using VG_328 as primary detection reagent. Different amounts (1.25–10 µl) of lysates of *E. coli* cultures (normalized by OD_{600}) expressing pD_NT and pD_NT_16 were separated by SDS-PAGE and subsequently blotted on a PVDF membrane. pD_NT contains the cognate peptide epitope; pD_NT_16 serves as negative control. VG_328 itself was detected via an anti-$RGSH_6$ antibody and a secondary goat anti-mouse IgG horseradish peroxidase conjugate. The arrow indicates pD_NT or pD_NT_16. On the right-hand side, a Coomassie-stained gel of 10 l of the lysates expressing pD_NT or pD_NT_16 is shown. (b) Pull-down experiment of binder VG_328 from *E. coli* lysate using the cognate peptide. Streptavidin-coated magnetic beads were incubated with biotinylated pD peptide fusions (pD_NT or pD_NT_16) or PBS-TB only. The beads were washed and subsequently incubated with a crude extract (CE) of *E. coli* expressing VG_328 (VG_328 CE) or PBS-TB only. After five washing steps, the beads were boiled in SDS loading buffer and the supernatant was loaded onto a 15% polyacrylamide gel, which was stained with Coomassie solution. Biotin carboxyl carrier protein (16.7 kDa) is found where no biotinylated protein has been immobilized on the streptavidin beads. In the last lane, 2 l of the CE was loaded for comparison.

monomeric in solution. We have optimized the ribosome display procedure and could select specific binders against the peptide target NT. The selected ArmRP allows residue-specific discrimination of peptide variants with four key side chains contributing most of the binding energy. The affinity is still moderate, and it will have to be determined whether binding is taking place in the canonical orientation

observed in natural ArmRPs. Nonetheless, the principal setup for new peptide selection technologies using designed ArmRP libraries has been established. Thus, the iterative further development of ArmRP protein libraries, alternating between protein design, computational approaches and structural verification, may allow us to progress toward the creation of a modular binding system for peptides.

Materials and Methods

General molecular biology methods

Unless stated otherwise, experiments were performed according to Sambrook and Russell.[47] Vent polymerase (New England Biolabs, Massachusetts, USA) was used for all DNA amplifications. Enzymes and buffers were from New England Biolabs or Fermentas (Lithuania). The cloning and production strain was *E. coli* XL1-blue (Stratagene, California, USA). Competent cells were prepared according to Inoue *Et al.*[48] Vector NTI (Invitrogen) was used for vector and oligonucleotide design. The cloning and protein expression vectors were pQE30 (Qiagen, Switzerland) and pPANK, a pQE30 derivative lacking the BpiI and BsaI sites (GenBank accession number AY327140). From this, the vector pPANK-NyCa was constructed by cloning of the capping repeats Ny and Ca.[29] pPANK-NyCa contains the BsaI and BpiI restriction sites between the capping repeats for cloning purposes. Note that the expression cassettes were constructed with a double stop codon.[29] pPANK-NyCa and pPANK-NyMMCa were used to clone the internal repeats for N3C and N5C library members. pPANK carries an $MRGSH_6$ tag at the N-terminus of the proteins. The DNA sequences corresponding to the NT and NT_A_4V_4 peptides were inserted into the vector pAT223 (GenBank accession number AY327138) and expressed as fusion proteins with pD to result in pD_NT and pD_NT_A_4V_4. The produced proteins consist of N-terminal Avi tag, pD, His_6 tag and the peptide of interest at the C-terminus (Supplementary Data Fig. S7b). The plasmid pBirAcm (Avidity, Colorado, USA), encoding *E. coli* biotin protein ligase BirA, was used for *in vivo* biotinylation of pD peptides.

Synthesis of ArmR modules

A complete list of all oligonucleotides used is given in Supplementary Data Table ST1. The trinucleotide phosphoramidites were obtained from Glen Research (USA), and lib5F was synthesized by Metabion (Germany). Standard oligonucleotides were from Microsynth (Switzerland).

An approach similar to the one described previously[29,32] was adopted for gene assembly (Supplementary Data Fig. S4). The details are given in Supplementary Methods. All single library modules were assembled by combining oligonucleotides indicated in Supplementary Data Table ST1 using assembly PCR. Due to randomization at position 41 (Fig. 1 and Supplementary Data Fig. S2), the end points of the capping repeats and the DNA sequence of the internal modules were therefore shifted by six bases compared to the original module M.[29]

The module sequence is shown in Supplementary Data Fig. S2. The designed ArmR library modules were PCR-assembled and subcloned for sequence analysis and to provide a template for the assembly of the whole library. In total, 2.2×10^7 independent clones for a single library module were obtained, a 20-fold oversampling of the theoretical repeat module diversity.

Synthesis of DNA encoding designed ArmRP library

The cells harboring the module library plasmids were re-grown in liquid culture to obtain sufficient DNA, and the plasmids were purified. This DNA was the starting material for the designed ArmRP library, which was assembled without additional frameshift selection. The single modules were PCR-amplified from the vectors, using external primers pQE_f_1 and pQE_r_1 (Qiagen). Modules were digested with the type IIS restriction enzymes BpiI and BsaI and directly ligated together (Supplementary Data Figs. S2, S4 and S5) essentially as described previously,[29] leading to a final library containing three internal randomized modules (N3C), as detailed in Supplementary Methods.

After each ligation step, the concentration of purified ligation product was measured. A ligation mix of sufficient DNA amount to represent and oversample the diversity of the designed ArmRP library was tested by transformation and sequencing of randomly picked clones. From the amount of DNA used after each ligation step, the practical library size of the N3C library was estimated at approximately 10^{11}. BamHI and KpnI restriction sites were used for insertion of the whole genes into the vector pPANK to obtain an initial N3C library.

The final N5C libraries were then formed by ligating double-digested PCR product of three library modules (LLL) into a plasmid already containing capping (N and C) and consensus (M) repeats, resulting in proteins with NMLLLMC module orientation. The full-length proteins were obtained as four types, characterized by the residues at positions 26 and 29 in each repeat (KK, KQ, QK or QQ).

Since the N-terminal cap from yeast (termed Ny or Y) and the artificial C-terminal cap (termed Ca or A) were to be used for many different constructs, a plasmid containing the two caps was also constructed (pPANK-NyCa). In this construct, the capping repeats are separated by a linker, containing the restriction sites for BpiI and BsaI. This construct made it possible to insert any internal repeat module directly in between the two caps (Supplementary Data Fig. S5).

Assembly of pD peptide fusions

Forward and reverse oligonucleotides (Supplementary Data Table ST1) encoding the peptide sequences and restriction sites for BamHI and HindIII were mixed and heated to 95 °C for 10 min and cooled to 4 °C to allow formation of double-stranded DNA. The double-stranded DNA fragments were then digested with BamHI and HindIII and inserted into the plasmid pAT223 (GenBank accession number AY327138).

Ribosome display

The libraries were originally ligated in vitro into the ribosome display vector pRDV (GenBank accession number AY327136) or its variant pRDVhis, containing an N-terminal $MRGSH_6$ tag instead of a FLAG tag. In the latter vector, a low in vitro translation rate was observed, which was unexpected, since it has been used robustly for ribosome display of Ankyrin repeat proteins (note that the vector with $MRGSH_6$ tag has been called there "pRDV" for short).[40,41] Upon further analysis, it was found that the low in vitro translation rate observed was most likely due to the formation of RNA secondary structure involving the starting ATG codon and a conserved part of the ArmR sequence, as hinted by the program M-fold[49–51] implemented in the GCG Package (Version 11.1.2; Accelrys Inc., California, USA). This hairpin is predicted to form by pairing with the sequence GTCCTCTC (nucleotide 79 in Supplementary Data Fig. S2; amino acids 21–23 in Fig. 1a), which is part of all internal repeats. Three silent mutations were introduced to disrupt the RNA secondary structure and brought the translation rate back to the expected level. The new pRDVhis_CAG (Supplementary Data Fig. S12) was used for the selections using N5C library.

To assemble the final N5C library, we cloned the three internal library modules (7.2 µg of insert) into the vector containing the NM–MC modules in the vector pPANK (4.5 µg). In preparation for ribosome display, plasmid pRDVhis_CAG was digested with BamHI and HindIII and purified from agarose gel using spin columns (QIAquick, Qiagen). We ligated 2 µg of plasmid fragment to 5 µg of insert DNA (about 10-fold molar excess), which had been previously digested with BamHI and HindIII. The ligation mix was amplified by PCR, and a practical library diversity $>10^{11}$ estimated from comparison to standards and considering the library quality was obtained. Forward primer T7B in combination with reverse primer tolA_Kurz was used for the large-scale PCR amplification of the designed ArmR libraries.[40,41] DNA sequencing was used to verify the correct sequences of some sample clones. A PCR product using outer primers T7b and tolAk (Supplementary Data Table ST1) provided the template for in vitro transcription.

The PCR-amplified N5C designed ArmRP DNA library was transcribed in vitro, and selection was performed by ribosome display essentially as described for DARPin libraries.[40,41]

For selection on plates, MaxiSorp plates (Nunc) were coated with NeutrAvidin (100 µl, 66 nM, overnight at 4 °C), blocked with BSA (300 µl, 0.5%, 1 h at room temperature), and target peptide (pD fusion or synthetic form of NT peptide) was immobilized via its biotin residue on NeutrAvidin [100 µl, 200 nM, 1 h at 4 °C; 10 mM Tris–HCl (pH 7.5) and 150 mM NaCl]. Binding and washing buffers were 50 mM Tris–HOAc (pH 7.5), 150 mM NaCl, 50 mM $Mg(OAc)_2$ and 0.5% BSA. The translation mix, containing the ternary mRNA–ribosome–designed ArmRPs complexes, was first prepanned in two wells (30 and 60 min) against biotinylated pD_NT_A_4V_4 or peptides with other sequence and immobilized the same way as NT or pD_NT to remove designed ArmRPs bound unspecifically. Subsequently, the translation mix was transferred to the well containing immobilized NT or pD_NT. The library was incubated for 45 min, and the washing time was increased from round to round (15 min total washing time in the first round to 75 min total washing time in the fourth round). After washing, the mRNA was eluted with 100 µl elution buffer [50 mM Tris–HOAc (pH 7.5), 150 mM NaCl and 25 mM ethylenediaminetetraacetic acid].

For selection on beads, the translation mix containing mRNA–ribosome–designed Armadillo complexes was incubated for 1 h at 4 °C with biotinylated target peptide in solution with varying concentrations in successive rounds. The complexes were captured by incubating with 50 µl streptavidin-coated paramagnetic beads (Dynabeads-MyOne Streptavidin) for 10 min at 4 °C. After washing the beads with WBT [50 mM Tris–HOAc (pH 7.6), 150 mM NaCl, 50 mM $Mg(OAc)_2$ and 0.01% Tween-20], the mRNA was eluted with elution buffer [50 mM Tris–HOAc (pH 7.6), 150 mM NaCl and 250 mM ethylenediaminetetraacetic acid] and prepared for another round of selection. To minimize selection of unspecific binders, we pretreated all tubes with TBST [50 mM Tris–HOAc (pH 7.6), 150 mM NaCl and 0.05% Tween-20] supplemented with 0.1% of BSA. To further avoid binders against the streptavidin present on the paramagnetic beads before each panning step, we pre-incubated the translation mix with the beads for 1 h at 4 °C and transferred the supernatant to a fresh tube.

Four rounds of ribosome display were performed, alternating targets and conditions as summarized in Table 1. The number of PCR cycles after RT was reduced from round to round from 45 to 35 to 30 to 25, adjusting to the yield due to progressive enrichment of binders in each round.

Subcloning and expression of ArmRPs

From the selected DNA pools showing specific binding to the respective target peptide, the DNA fragments encoding the designed ArmRP inserts were amplified by PCR and subcloned into the expression vector pQE30ss (containing a double stop codon TAA-TGA) via BamHI/KpnI using oligonucleotides: Ny1F and Ca6R (Supplementary Data Table ST1). For large-scale expression, *E. coli* XL1-blue cells were transformed with the respective plasmid and grown in tryptone–yeast extract–NaCl medium containing 1% (w/v) glucose and 50 g/ml ampicillin at 37 ℃ with vigorous shaking. Expression was induced by IPTG (final concentration of 0.5 mM) when the culture reached $OD_{600} = 0.6$. After 4 h of expression, cells were harvested by centrifugation.

Protein expression, purification and characterization

pD fusion proteins were expressed in *E. coli* XL1-blue cells as described for ArmRPs. For *in vivo* biotinylation of pD peptide fusions that contain an N-terminal Avi tag (GLNDIEAQKIEWHE), cells were co-transformed with pBirAcm and pAT223 (carrying the pD peptide fusion constructs) and grown at 37 ℃ in medium containing 30 g/ml chloramphenicol and 50 g/ml ampicillin. Before induction with IPTG, biotin was added to the medium to a final concentration of 50 M, according to Cull and Schatz.[52] All further steps were carried out at 4 ℃. After 4 h, cells were resuspended in 50 mM Tris–HCl and 500 mM NaCl (pH 8.0) and lysed in a French pressure cell (SLM Instruments, New York, USA) at a pressure of 1200 psi. The lysis mixture was further homogenized by sonication (Branson, Missouri, USA). Insoluble material was pelleted

by centrifugation at 20,000***g*** for 30 min. The supernatant was purified by IMAC with Ni-NTA agarose (Qiagen) and equilibrated with buffer containing 50 mM Tris–HCl, 500 mM NaCl, 10% (v/v) glycerol and 20 mM imidazole (pH 8.0). Columns were washed extensively with the equilibration buffer and then proteins were eluted with an elution buffer identical with the equilibration buffer but additionally containing 250 mM imidazole. Samples were then dialyzed overnight against 50 mM Tris–HCl, pH 8.5, and applied to a Mono Q anion-exchange column (GE Healthcare) equilibrated with running buffer (50 mM Tris–HCl, pH 8.5) using the ÄKTA explorer chromatography system. The column was then washed with 50 mM Tris–HCl and 20 mM NaCl (pH 8.5), and the samples eluted with a gradient from 20 mM to 1 M NaCl.

These constructs were further analyzed by mass spectrometry for their homogenous preparation and extent of biotinylation, which was determined as close to 100%. Protein concentrations were determined by absorbances at 235 and 280 nm using extinction coefficients calculated with the tools available at the ExPASy proteomics server† and by the bicinchoninic acid assay (Pierce).

SEC and MALS

Analytical SEC was carried out on an ÄKTA explorer chromatography system using a Superdex 200 10/30 GL column ($V_0 \approx 8$ ml, $V_{tot} \approx 24$ ml and flow rate = 0.5 ml/min) (GE Healthcare). Phosphate buffer (50 mM phosphate and 150 mM NaCl, pH 7.4) and Tris-based buffers (20 mM Tris–HCl and 50 mM NaCl, pH 8.0) were used. The random clones and selected designed ArmRPs were analyzed in buffer containing 20 mM Tris–HCl and 50 mM NaCl, pH 8.0. SEC–MALS measurements were performed on an Agilent 1100 HPLC system or an ÄKTA system connected to a triangle light-scattering detector and a differential refractometer (miniDAWN Tristar and Optilab, respectively; Wyatt Technology, California, USA). With a well plate autosampler (Agilent 1100), sample volumes of 75 µl with a minimum protein concentration of 30 µM were loaded on a Superdex 200 10/300 GL column (GE Healthcare) with a flow rate of 0.5 ml/min. A specific refractive index increment (dn/dc) value of 0.186 ml/g was used for the protein.[53] The data were recorded and processed using the ASTRA 4.73.04 or ASTRA V software (Wyatt Technology). To determine the detector delay volumes and normalization coefficients for the MALS detector, we used a BSA sample (Sigma USA, A8531) as reference. Neither despiking nor a band broadening correction was applied.

CD spectroscopy

CD measurements were performed on a Jasco J-810 spectropolarimeter (Jasco, Japan) using a 0.5-mm cylindrical thermocuvette. CD spectra were recorded from 190 to 250 nm with a data pitch of 1 nm, a scan speed of 20 nm/min, a response time of 4 s and a bandwidth of 1 nm. Each spectrum was recorded three times and averaged. Measurements were performed at 20 °C. The CD signal was corrected by buffer subtraction and converted to MRE (mean residue *E*llipticity). Heat denaturation curves were obtained by measuring the CD signal at 222 nm with the temperature increasing from 20 to 95 ℃ (data pitch, 1 nm; heating rate, 1 ℃/min; response time, 4 s; bandwidth, 1 nm). Data were processed as described above. Guanidinium-induced denaturation measurements were performed after overnight incubation at 20 ℃ with increasing concentrations of GdmCl (99.5% purity; Fluka, Switzerland), and the data were collected and processed as described above. Measurements of designed ArmRPs were performed in phosphate buffer (50 mM phosphate and 150 mM NaCl, pH 7.4).

ANS binding

ANS fluorescence was measured using a PTI QM-2000-7 fluorimeter (Photon Technology International, New Jersey, USA). The measurements were performed at 20 °C in 20 mM Tris–HCl, 50 mM NaCl and 100 µM ANS (pH 8.0) using purified proteins at a final concentration of 10 µM. The emission spectrum from 400 to 650 nm (1 nm/s) was recorded with an excitation wavelength of 350 nm. For each sample, three spectra were recorded and averaged.

ELISA

A MaxiSorp plate (Nunc) was coated with NeutrAvidin (100 µl, 66 nM, overnight at 4 °C) and then blocked with BSA (300 µl, 0.5%, 1 h at room temperature); the target peptide (pD fusion or synthetic form of NT) was immobilized via its biotin residue on NeutrAvidin [100 µl, 200 nM, in 100 mM phosphate (pH 7.0) and 150 mM NaCl]. ELISA buffers for binding and washing in all ELISA experiments were PBS-B [50 mM phosphate and 150 mM NaCl (pH 7.4) with 0.3% BSA] and PBS-BT [50 mM phosphate and 150 mM NaCl (pH 7.4) with 0.3% BSA and 0.01% Tween-20], if not stated otherwise.

For crude extract ELISAs, single clones were grown in 96-deep-well plates overnight at 37 °C. After induction for 4 h, cells were harvested, resuspended in 100 µl B-PERII (Pierce) and incubated for 15 min at room temperature with vortexing from time to time. Then, 900 µl PBS-BT was added and cell debris was removed by centrifugation.

For the ELISA, 10–100 µl of the above crude extracts was applied and incubated for 1 h at 4 °C. After extensive washing of the wells three times with 300 µl of 1× PBS-BT, we incubated wells with an anti-$RGSH_6$ antibody (1:5000 dilution in 1× PBS-BT, 1 h at 4 °C; Qiagen, Germany) as primary antibody. A goat anti-mouse IgG alkaline phosphatase conjugate (Sigma) (1:10,000 in 1× PBS-BT, 1 h at 4 °C) was used as secondary antibody with the substrate disodium 4-nitrophenyl phosphate (3 mM) (Fluka) in buffer containing 50 mM $NaHCO_3$ and 50 mM $MgCl_2$. Absorbance at 405 nm (540 nm reference wavelength) was measured using a Perkin Elmer HTS 7000 Plus plate reader.

For ELISAs with purified proteins, coating was carried out as above. The wells were blocked with 300 µl of 1× PBS-TB (PBS containing 0.1% Tween-20 and 0.3% BSA) for 1 h at room temperature. Immobilization of the biotinylated target proteins (100 µl of 40–100 nM for ELISA) in PBS-TB was allowed to occur for 1 h at 4 °C. Binding was detected as described above. All proteins were dissolved in PBS-B, and all washing steps were carried out in PBS-BT. Detection was carried out as described above.

For competition, purified designed ArmRPs were incubated with different concentrations of non-biotinylated target peptides or fusion proteins present before (1 h at

4 °C) and during (8 min at 4 °C) the binding reaction. Washing and further detection of binding was performed as detailed above.

Surface plasmon resonance

SPR was measured using a BIACORE 3000 instrument (GE Healthcare Biosciences, Pennsylvania, USA). The running buffer was PBS-T [50 mM phosphate and 150 mM NaCl (pH 7.4) with 0.01% Tween-20]. A streptavidin SA chip (GE Healthcare Biosciences) was used with 10 RU (response units) biotinylated synthetic NT_(1–13) immobilized. The interactions were measured at a flow rate of 30 µl/min with 2 min buffer flow, 2 min injection of VG_328 in varying concentrations (0.1–215 µM) and dissociation for 10 min with buffer flow. The signal of an uncoated reference cell was subtracted from the measurements. The dissociation constant was determined from the plateau values using Scrubber (BioLogic software).

Western blots

pD_NT and pD_NT_16 were expressed in *E. coli* XL1-blue cells as described for ArmRPs. Cells from a 1-ml aliquot of the expression culture were pelleted and resuspended in a volume of PBS corresponding to 500 l times the final OD_{600} of the culture. The resuspended cells (10 l, 5 l, 2.5 l or 1.25 l) were mixed with SDS loading buffer (20% final) and incubated at 96 ℃ for 15 min. Samples were loaded on a 15% acrylamide gel and run at 180 V. Subsequently, the gel was blotted on a PVDF (*p*olyvinyli*d*ene *f*luoride) membrane using a semidry blotting device (Biorad) and transfer buffer [20 mM Tris–HCl (pH 8.3), 150 mM glycine, 0.02% (w/v) SDS and 20% (v/v) methanol]. The membrane was blocked with PBS-T containing 3% BSA at 4 ℃ overnight. The membrane was washed twice in PBS-T. The blot was then incubated with 5 ml of VG_328 [40 M in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder] for 1 h at room temperature in a 50-ml Falcon tube on a roller mixer and subsequently washed three times for 5 min in PBS-T (20 ml each). Bound VG_328 was detected via an anti-$RGSH_6$ antibody [Qiagen, Germany; 1:5000 dilution in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder] for 1 h at room temperature, followed by subsequent washing three times for 5 min in PBS-T (20 ml each). A goat anti-mouse IgG horseradish peroxidase conjugate was used as secondary antibody [1:10,000 dilution in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder; Pierce] with subsequent washing five times for 10 min in PBS-T (20 ml each). The membrane was incubated with horseradish peroxidase substrate (Millipore Immobilon Western) for 1 min. A chemiluminescent picture was taken on a Fuji-Film LAS-3000 device.

Pull-down experiments

VG_328 was expressed in *E. coli* XL1-blue cells as described for ArmRPs. A crude extract was prepared by lysis of the cells with sonication (Branson), insoluble cell debris was pelleted (30 min, 20,000***g***) and the supernatant was filtered (0.22 m). Tubes and magnetic streptavidin beads (DynabeadsMyOne Streptavidin T1; Invitrogen) were blocked with PBS-TB (1 h, room temperature). Beads (20 l) were incubated with 10 l of purified, *in vivo* biotinylated pD_NT or pD_NT_16 (100 M) or PBS-TB (1 h at 4 ℃). The beads were washed twice with 500 l PBS-T and then incubated with 200 l crude extract of the VG_328 expression culture (1 h at 4 ℃). After five times washing with 500 l PBS-T, the beads were resuspended in 15 l PBS and 5 l SDS loading buffer was added. This mixture was incubated for 10 min at 96 ℃, the beads were pelleted (10 s, 10,000***g***) and the supernatant was loaded on a 15% acrylamide gel. The gel was run at 180 V and stained with Coomassie solution.

Acknowledgements

This work was supported by Sinergia grant CRSI33_122686 from the Schweizerischer Nationalfonds to A.P., by predoctoral fellowships from the Forschungskredit der Universität Zürich to G.V. and from the Roche Research Foundation to F.P. and by a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft to D.T. We thank Prof. Oliver Zerbe for critical reading of the manuscript.

Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2012.08.029

Received 15 April 2012;
Received in revised form 8 June 2012;
Accepted 23 August 2012
Available online 14 September 2012

Keywords:
combinatorial library;
Armadillo repeat;
ribosome display;
neurotensin;
peptide recognition

† http://web.expasy.org/protparam/

Present address: F. Parmeggiani, Department of Biochemistry, University of Washington, Health Sciences Building, Box 357350, Seattle, WA 98195, USA.

Abbreviations used:
ArmR, Armadillo repeat; ArmRP, Armadillo repeat protein; BSA, bovine serum albumin; IMAC, immobilized metal-ion affinity chromatography; MALS, multi-angle light scattering; SEC, size-exclusion chromatography; SPR, surface plasmon resonance; ANS, 8-anilino-napthalene-1-sulfonate; pD, protein D; NLS, nuclear localization sequence; NT, neurotensin.

References

1. Binz, H. K., Amstutz, P. & Plückthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **23**, 1257–1268.
2. Boersma, Y. L. & Plückthun, A. (2011). DARPins and other repeat protein scaffolds: advances in engineering and applications. *Curr. Opin. Biotechnol.* **22**, 849–857.
3. Löfblom, J., Frejd, F. Y. & Ståhl, S. (2011). Non-immunoglobulin based protein scaffolds. *Curr. Opin. Biotechnol.* **22**, 843–848.
4. Clonis, Y. D. (2005). Affinity chromatography matures as bioinformatic and combinatorial tools develop. *J. Chromatogr., A*, **1101**, 1–24.
5. Spisak, S. & Guttman, A. (2009). Biomedical applications of protein microarrays. *Curr. Med. Chem.* **16**, 2806–2815.
6. MacCallum, R. M., Martin, A. C. & Thornton, J. M. (1996). Antibody–antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.* **262**, 732–745.
7. Wilson, I. A., Ghiara, J. B. & Stanfield, R. L. (1994). Structure of anti-peptide antibody complexes. *Res. Immunol.* **145**, 73–78.
8. Almagro, J. C. (2004). Identification of differences in the specificity-determining residues of antibodies that recognize antigens of different size: implications for the rational design of antibody repertoires. *J. Mol. Recognit.* **17**, 132–143.
9. Malabarba, M. G., Milia, E., Faretta, M., Zamponi, R., Pelicci, P. G. & Di Fiore, P. P. (2001). A repertoire library that allows the selection of synthetic SH2s with altered binding specificities. *Oncogene*, **20**, 5186–5194.
10. Panni, S., Dente, L. & Cesareni, G. (2002). In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *J. Biol. Chem.* **277**, 21666–21674.
11. Hiipakka, M., Poikonen, K. & Saksela, K. (1999). SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. *J. Mol. Biol.* **293**, 1097–1106.
12. Dalby, P. A., Hoess, R. H. & DeGrado, W. F. (2000). Evolution of binding affinity in a WW domain probed by phage display. *Protein Sci.* **9**, 2366–2376.
13. Cowburn, D. (1997). Peptide recognition by PTB and PDZ domains. *Curr. Opin. Struct. Biol.* **7**, 835–838.
14. Schneider, S., Buchert, M., Georgiev, O., Catimel, B., Halford, M., Stacker, S. A. *Et al.* (1999). Mutagenesis and selection of PDZ domains that bind new protein targets. *Nat. Biotechnol.* **17**, 170–175.
15. Ernst, A., Gfeller, D., Kan, Z., Seshagiri, S., Kim, P. M., Bader, G. D. & Sidhu, S. S. (2010). Coevolution of PDZ domain–ligand interactions analyzed by high-throughput phage display and deep sequencing. *Mol. Biosyst.* **6**, 1782–1790.
16. Esteban, O. & Zhao, H. (2004). Directed evolution of soluble single-chain human class II MHC molecules. *J. Mol. Biol.* **340**, 81–95.
17. Grove, T. Z., Cortajarena, A. L. & Regan, L. (2008). Ligand binding by repeat proteins: natural and designed. *Curr. Opin. Struct. Biol.* **18**, 507–515.
18. Kajava, A. V. (2001). Review: proteins with repeated sequence—structural prediction and modeling. *J. Struct. Biol.* **134**, 132–144.
19. Andrade, M. A., Petosa, C., O'Donoghue, S. I., Muller, C. W. & Bork, P. (2001). Comparison of ARM and HEAT protein repeats. *J. Mol. Biol.* **309**, 1–18.
20. Marfori, M., Mynott, A., Ellis, J. J., Mehdi, A. M., Saunders, N. F., Curmi, P. M. *Et al.* (2011). Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim. Biophys. Acta*, **1813**, 1562–1577.
21. Xu, W. & Kimelman, D. (2007). Mechanistic insights from structural studies of β-catenin and its binding partners. *J. Cell Sci.* **120**, 3337–3344.
22. Tewari, R., Bailes, E., Bunting, K. A. & Coates, J. C. (2010). Armadillo-repeat protein functions: questions for little creatures. *Trends Cell Biol.* **20**, 470–481.
23. Hatzfeld, M. (1999). The armadillo family of structural proteins. *Int. Rev. Cytol.* **186**, 179–224.
24. Cortajarena, A. L. & Regan, L. (2006). Ligand binding by TPR domains. *Protein Sci.* **15**, 1193–1198.
25. Conti, E., Uy, M., Leighton, L., Blobel, G. & Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin α. *Cell*, **94**, 193–204.
26. Conti, E. & Kuriyan, J. (2000). Crystallographic analysis of the specific yet versatile recognition of distinct nuclear localization signals by karyopherin α. *Structure*, **8**, 329–338.
27. Huber, A. H. & Weis, W. I. (2001). The structure of the β-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by β-catenin. *Cell*, **105**, 391–402.
28. Madhurantakam, C., Varadamsetty, G., Grütter, M. G., Plückthun, A. & Mittl, P. R. (2012). Structure-based optimization of designed Armadillo-repeat proteins. *Protein Sci.* **21**, 1015–1028.
29. Parmeggiani, F., Pellarin, R., Larsen, A. P., Varadamsetty, G., Stumpp, M. T., Zerbe, O. *Et al.* (2008). Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core. *J. Mol. Biol.* **376**, 1282–1304.
30. Catimel, B., Teh, T., Fontes, M. R., Jennings, I. G., Jans, D. A., Howlett, G. J. *Et al.* (2001). Biophysical characterization of interactions involving importin- during nuclear import. *J. Biol. Chem.* **276**, 34189–34198.
31. 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.
32. 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.
33. Plückthun, A. (2012). Ribosome display: a perspective. *Methods Mol. Biol.* **805**, 3–28.
34. Alfarano, P., Varadamsetty, G., Ewald, C., Parmeggiani, F., Pellarin, R., Zerbe, O. *Et al.* (1996). Optimization of designed armadillo repeat proteins by molecular dynamics simulations and NMR spectroscopy. *Protein Sci.* **21**, 1298–1314.
35. Tripp, K. W. & Barrick, D. (2007). Enhancing the stability and folding rate of a repeat protein through the

addition of consensus repeats. *J. Mol. Biol.* **365**, 1187–1200.
36. Kajander, T., Cortajarena, A. L., Main, E. R., Mochrie, S. G. & Regan, L. (2005). A new folding paradigm for repeat proteins. *J. Am. Chem. Soc.* **127**, 10188–10190.
37. Wetzel, S. K., Settanni, G., Kenig, M., Binz, H. K. & Plückthun, A. (2008). Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. *J. Mol. Biol.* **376**, 241–257.
38. Nieto, J. L., Rico, M., Santoro, J., Herranz, J. & Bermejo, F. J. (1986). Assignment and conformation of neurotensin in aqueous solution by ^{1}H NMR. *Int. J. Pept. Protein Res.* **28**, 315–323.
39. Hanes, J. & Plückthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. USA*, **94**, 4937–4942.
40. 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.
41. Dreier, B. & Plückthun, A. (2012). Rapid selection of high-affinity binders using ribosome display. *Methods Mol. Biol.* **805**, 261–286.
42. Binz, H. K., Amstutz, P., Kohl, A., Stumpp, M. T., Briand, C., Forrer, P. *Et al.* (2004). High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* **22**, 575–582.
43. Hodel, A. E., Harreman, M. T., Pulliam, K. F., Harben, M. E., Holmes, J. S., Hodel, M. R. *Et al.* (2006). Nuclear localization signal receptor affinity correlates with *in vivo* localization in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **281**, 23545–23556.
44. Huang, J., Makabe, K., Biancalana, M., Koide, A. & Koide, S. (2009). Structural basis for exquisite specificity of affinity clamps, synthetic binding proteins generated through directed domain-interface evolution. *J. Mol. Biol.* **392**, 1221–1231.
45. Li, S. J. & Cronan, J. E., Jr (1992). The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase. *J. Biol. Chem.* **267**, 855–863.
46. Fall, R. R., Alberts, A. W. & Vagelos, P. R. (1975). Analysis of bacterial biotin-proteins. *Biochim. Biophys. Acta*, **379**, 496–503.
47. Sambrook, J. & Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual, 3rd edit Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
48. Inoue, H., Nojima, H. & Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, **96**, 23–28.
49. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
50. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989). Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. USA*, **86**, 7706–7710.
51. Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.
52. Cull, M. G. & Schatz, P. J. (2000). Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol.* **326**, 430–440.
53. Wen, J., Arakawa, T. & Philo, J. S. (1996). Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal. Biochem.* **240**, 155–166.

jmb Erratum

Erratum to "Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High Specificity" [*J. Mol. Biol.* (2012) 424, 68–87]

Gautham Varadamsetty, Dirk Tremmel, Simon Hansen, Fabio Parmeggiani and Andreas Plückthun

Department of Biochemistry, *University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

Correspondence to Andreas Plückthun: *plueckthun@bioc.uzh.ch*
DOI of original article: http://dx.doi.org/10.1016/j.jmb.2012.08.029
http://dx.doi.org/10.1016/j.jmb.2013.01.012

Due to a typesetting error, the above-referenced article was published with errors. In several instances, the "micro" sign (μ, greek mu), which was entirely correct in the online "Corrected Proof", completely disappeared in the printed and online version of the article.

On page 81, the legend for Fig. 9 should read:

Fig. 9. (a) Western blot (left) using VG_328 as primary detection reagent. Different amounts (1.25–10 μl) of lysates of *E. coli* cultures (normalized by OD_{600}) expressing pD_NT and pD_NT_16 were separated by SDS-PAGE and subsequently blotted on a PVDF membrane. pD_NT contains the cognate peptide epitope; pD_NT_16 serves as negative control. VG_328 itself was detected via an anti-$RGSH_6$ antibody and a secondary goat anti-mouse IgG horseradish peroxidase conjugate. The arrow indicates pD_NT or pD_NT_16. On the right-hand side, a Coomassie-stained gel of 10 μl of the lysates expressing pD_NT or pD_NT_16 is shown. (b) Pull-down experiment of binder VG_328 from *E. coli* lysate using the cognate peptide. Streptavidin-coated magnetic beads were incubated with biotinylated pD peptide fusions (pD_NT or pD_NT_16) or PBS-TB only. The beads were washed and subsequently incubated with a crude extract (CE) of *E. coli* expressing VG_328 (VG_328 CE) or PBS-TB only. After five washing steps, the beads were boiled in SDS loading buffer and the supernatant was loaded onto a 15% polyacrylamide gel, which was stained with Coomassie solution. Biotin carboxyl carrier protein (16.7 kDa) is found where no biotinylated protein has been immobilized on the streptavidin beads. In the last lane, 2 μl of the CE was loaded for comparison.

On page 85, the sections "Western blots" and "Pull-down experiments" should read:

Western blots

pD_NT and pD_NT_16 were expressed in *E. coli* XL1-blue cells as described for ArmRPs. Cells from a 1-ml aliquot of the expression culture were pelleted and resuspended in a volume of PBS corresponding to 500 μl times the final OD_{600} of the culture. The resuspended cells (10 μl, 5 μl, 2.5 μl or 1.25 μl) were mixed with SDS loading buffer (20% final) and incubated at 96 °C for 15 min. Samples were loaded on a 15% acrylamide gel and run at 180 V. Subsequently, the gel was blotted on a PVDF (polyvinylidene fluoride) membrane using a semidry blotting device (Biorad) and transfer buffer [20 mM Tris–HCl (pH 8.3), 150 mM glycine, 0.02% (w/v) SDS and 20% (v/v) methanol]. The membrane was blocked with PBS-T containing 3% BSA at 4 °C overnight. The membrane was washed twice in PBS-T. The blot was then incubated with 5 ml of VG_328 [40 μM in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder] for 1 h at room temperature in a 50-ml Falcon tube on a roller mixer and subsequently washed three times for 5 min in PBS-T (20 ml each). Bound VG_328 was detected via an anti-$RGSH_6$ antibody [Qiagen, Germany; 1:5000 dilution in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder] for 1 h at room temperature, followed by subsequent washing three times for 5 min in PBS-T (20 ml each). A goat anti-mouse IgG horseradish peroxidase conjugate was used as secondary

Please cite this article as: Varadamsetty, G. *et al.*, Erratum to "Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High..., *J. Mol. Biol.* (2013), http://dx.doi.org/10.1016/j.jmb.2013.01.012

antibody [1:10,000 dilution in PBS-T containing 1% (w/v) BSA and 2% (w/v) milk powder; Pierce] with subsequent washing five times for 10 min in PBS-T (20 ml each). The membrane was incubated with horseradish peroxidase substrate (Millipore Immobilon Western) for 1 min. A chemiluminescent picture was taken on a Fuji-Film LAS-3000 device.

Pull-down experiments

VG_328 was expressed in *E. coli* XL1-blue cells as described for ArmRPs. A crude extract was prepared by lysis of the cells with sonication (Branson), insoluble cell debris was pelleted (30 min, 20,000*g*) and the supernatant was filtered (0.22 μm). Tubes and magnetic streptavidin beads (DynabeadsMyOne Streptavidin T1; Invitrogen) were blocked with PBS-TB (1 h, room temperature). Beads (20 μl) were incubated with 10 μl of purified, *in vivo* biotinylated pD_NT or pD_NT_16 (100 μM) or PBS-TB (1 h at 4 °C). The beads were washed twice with 500 μl of PBS-T and then incubated with 200 μl of crude extract of the VG_328 expression culture (1 h at 4 °C). After five times washing with 500 μl of PBS-T, the beads were resuspended in 15 μl of PBS and 5 μl of SDS loading buffer was added. This mixture was incubated for 10 min at 96 °C, the beads were pelleted (10 s, 10,000*g*) and the supernatant was loaded on a 15% acrylamide gel. The gel was run at 180 V and stained with Coomassie solution.

The publisher regrets any confusion caused by these errors.

Present address: F. Parmeggiani, Department of Biochemistry,
University of Washington, Health Sciences Building,
Box 357350, Seattle, WA 98195, USA.

Please cite this article as: Varadamsetty, G. *et al.*, Erratum to "Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High..., *J. Mol. Biol.* (2013), http://dx.doi.org/10.1016/j.jmb.2013.01.012