LETTER

ACS chemical biology

Designed Proteins To Modulate Cellular Networks

Aitziber L. Cortajarena[†,§], Tina Y. Liu[†,§], Mark Hochstrasser[†], and Lynne Regan[†,‡,*]

[†]Department of Molecular Biophysics & Biochemistry and [‡]Department of Chemistry, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06520, [§]These authors contributed equally to this work.

ABSTRACT A major challenge of protein design is to create useful new proteins that interact specifically with biological targets in living cells. Such binding modules have many potential applications, including the targeted perturbation of protein networks. As a general approach to create such modules, we designed a library with approximately 10^9 different binding specificities based on a small 3-tetratricopeptide repeat (TPR) motif framework. We employed a novel strategy, based on split GFP reassembly, to screen the library for modules with the desired binding specificity. Using this approach, we identified modules that bind tightly and specifically to Dss1, a small human protein that interacts with the tumor suppressor protein BRCA2. We showed that these modules also bind the yeast homologue of Dss1, Sem1. Furthermore, we demonstrated that these modules inhibit Sem1 activity in yeast. This strategy will be generally applicable to make novel genetically encoded tools for systems/synthetic biology applications.

Proteins perform a vast array of different cellular activities, but none function in isolation. Instead, they act synergistically as a system of complex networks of exquisitely regulated protein−protein interactions (*1–6*). Genome-scale analyses have begun to reveal the components of protein interaction networks and their role in specific cellular processes, pathways, and disease. One can now envisage specific interventions to inhibit pathways and consequently change cellular and even organismal fate (*7, 8*). In order to specifically modulate interactions *in vivo*, there is a need for new molecular tools. Recently, there have been several efforts to develop engineered protein scaffolds that can be used as intracellular binding module alternatives of coventional antibodies, covered in this review (*9*). Here we present a novel strategy by which to create small protein modules to bind to any desired protein target, which can be used to perturb protein interaction networks. We demonstrate that it is possible to create protein modules (T-Mods) tailored to have desired binding specificities and affinities and to use them as tools to modulate cellular processes *in vivo*. We show the application of such a T-Mod to inhibit an essential protein−protein interaction in yeast. The principles and approach taken are general and should prove a useful tool applicable to many other systems.

Our approach is to generate a library of binding modules, based on a small, stable protein framework, and to screen this library to identify modules with the desired binding specificity. Having identified modules with the desired binding specificity, we expressed them *in vivo* to inhibit a specific protein−protein interaction, and analyzed their effects on cellular pathways. The tailored binding modules (T-Mods) are based on the tetratricopeptide repeat (TPR) motif, a 34-amino-acid sequence that adopts a helix-turn-helix structure (Figure 1, panel a) (*10–12*). The TPR helices stack to form elongated superhelical structures, with the inner surface of the superhelix serving as a platform for protein−protein interactions (*10, 13–18*). Three tandem repeats comprise both the smallest functional array of TPR motifs and the most common number of repeats per domain in nature (*10*). Proteins containing TPR domains display a broad range of binding specificities and are implicated in a variety of cellular processes. In addition, TPR domains bind their ligands without undergoing major conformational changes (*19*) and are stable and robust scaffolds (*16, 18*). All these features make the TPR fold an attractive structural framework for the design of novel binding modules (*20–24*).

We chose a biologically and clinically relevant protein as a test case: the human protein Dss1. Dss1 is a small, evolutionarily conserved protein that was first identified through the association of its deficient expression with the split-hand/split-foot malformation disorder (*25*). Dss1 is a binding

*Corresponding author, lynne.regan@yale.edu.

Received for review July 17, 2009 and accepted December 19, 2009.

Published online December 20, 2009

10.1021/cb9002464

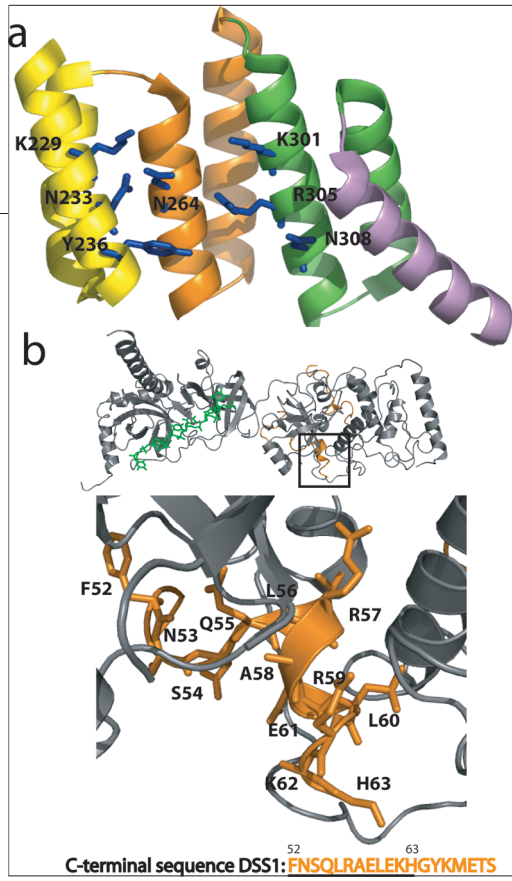

Figure 1. TPR framework and Dss1 structures. a) Structure of the TPR2A domain, used as our scaffold, where each repeat and the solvating helix are colored differently. The seven residues that we designated for randomization are highlighted in blue. All of the residues selected for randomization (K229, N233, Y236, N264, K301, R305, and N308 of Hsp Organizing protein, Hop, where the TPR2A domain is amino acids 222−352) interact with Hsp90s C-terminal peptide. Five of these residues, known as a "carboxylate clamp" (K229, N233, N264, K301, and R305), are critical for TPR2A's ability to bind Hsp90 (*13, 15*). b) Mammalian breast cancer type 2 (BRCA2) protein's DNA-binding domain complexed with ssDNA and Dss1 (*30*). Dss1 is shown in ribbon representation, highlighted in orange. BRCA2 is shown in grey and the DNA is shown in green. The figure below shows a close-up of the area highlighted in the black box, showing the C-terminal residues of the Dss1 protein that were targeted in the selection. Dss1-C19 19-mer peptide sequence, the 12 underlined residues, are seen in the X-ray structure interacting with BRCA2; the last 7 C-terminal residues are not seen in the structure.

partner of the breast cancer susceptibility protein, BRCA2 (*26*). BRCA2 is involved in double-strand break DNA repair (*27*), and Dss1 binding is critical for both BRCA2 stability and function (*28, 29*). The residues of both Dss1 and BRCA2 that form the interaction interface are particularly well conserved across different species relative to the rest of the protein. Also, mutations at several of the residues of BRCA2 that participate in Dss1 binding are associated with breast cancer (*30*) (Figure 1, panel b). Tantalizingly, although yeast do not possess a BRCA2 homologue, they do have a Dss1 ortholog, called Sem1, which resides primarily in the nucleus. Cellular roles that have been proposed for Sem1 include mRNA export (*31, 32*), proteasomal function (*33–35*), and RNA splicing (*32*). The specific function of Sem1 in any of these pathways, however, remains unclear. By designing a protein that can bind Dss1 or Sem1 *in vivo*, we generated a valuable tool with which to elucidate and modulate the mechanism of action of this protein.

To identify modules with the desired binding specificity, we first created a library of 2.7×10^8 different proteins using the TPR2A domain of Hop (*13*) as a scaffold, by simultaneously randomizing seven key residues in the binding site to each of 19 different amino acids (Figure 1, panel a) (*36, 37*). To identify the TPR proteins that bind Dss1, we used a split green fluorescent protein (GFP)-based screen (*38, 39*), a new method of screening libraries to identify binding partners (see the following Article). We have previously demonstrated that the two halves of GFP will only reassemble *in vivo* to generate a fluorescent protein when they are expressed as fusions to two proteins that interact. Most importantly, the split GFP assay recapitulates binding specificities that have been measured *in vitro,* and the possibility of false positives is therefore extremely low (*38, 39*). We fused the C-terminal 19 residues of Dss1 (Dss1-C19) (Figure 1, panel b) to the N-terminal fragment of GFP (N-GFP) and the TPR library to the C-terminal fragment of GFP (C-GFP). The last seven residues of Dss1 are not observed in the X-ray structure and are presumably unstructured; the preceding 12 residues of Dss1 interact with BRCA2 (Figure 1, panel b) (*30*).

We used three rounds of fluorescence-activated cell sorting (FACS) to identify and collect *Escherichia coli* cells with high fluorescence (see the following Article). Fluorescent cells from the final sort were collected, plated, and restreaked to confirm fluorescence, and 96 of 106 colonies were fluorescent upon restreaking. From these 96 hits, 84 complete sequences were obtained and of these 55 were unique.

An examination of amino acid identity at the randomized positions of the 55 unique genes clearly showed that certain amino acids or types of amino acids are favored at three positions (K301, R305, and N308, Figure 1, panel a). An acidic residue (Glu or Asp) was favored at position 301 in 80% of the hits. At position 305 the two most common residues were Leu and Val (67%) and at position 308 Leu and Met were most common (80%). The strong consensus at these positions suggested that the selection had worked, and that these residues may constitute the binding site for Dss1-C19. No obvious consensus was observed at the other four randomized positions.

We chose two T-Mods for further characterization: T-mod(Dss1A), a consensus hit with Asp, Leu, Leu at positions 301, 305, and 308 respectively, and T-mod(Dss1B), a nonconsensus hit with Glu, Val, and Arg at those three positions. The genes for the chosen TPR modules were subcloned into an expression vector, and the proteins purified (*16, 21*). Both are folded and of comparable stability as the parent TPR, as determined by circular dichroism (data not shown) (*19*).

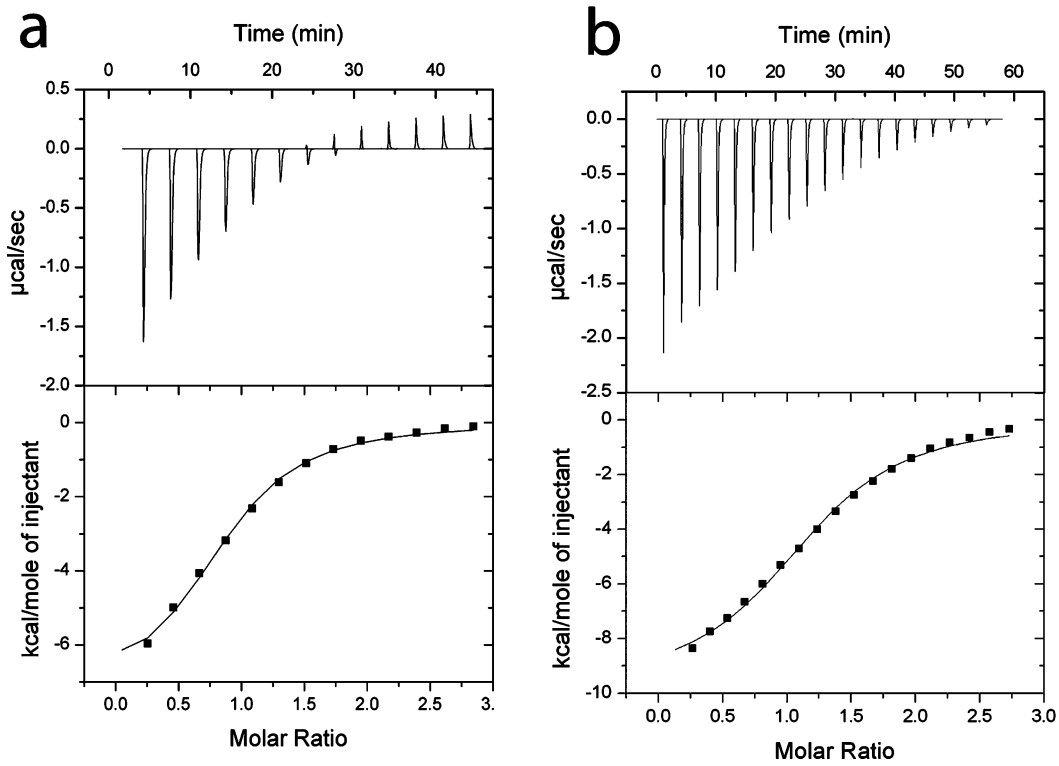

Figure 2. T-Mods *in vitro* binding by ITC. ITC binding isotherms for the interaction of Dss1 19-mer C-terminal peptide with two Dss1-binding T-Mods: T-Mod(Dss1A) (a) and T-Mod(Dss1B) (b). Dss1-C19 peptide was titrated at 2.66 mM into 130 μM T-Mod(Dss1A) solution, and 1.52 mM Dss1-C19 was titrated in 114 μM T-Mod(Dss1B) solution. The data was fit to a 1:1 binding model to calculate the stoichiometry (*N*) and the binding affinity (K_d) of the interactions using Origin 7.0 (T-Mod(Dss1A) K_d = 20.2 μM, N = 0.8; T-Mod(Dss1B) K_d = 18.6 μM, N = 1.1).

We determined the binding affinity of the selected T-Mods for Dss1 using isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR). T-Mod(Dss1A) binds Dss1-C19 with a dissociation constant of 20 μM and T-Mod(Dss1B) binds Dss1-C19 with a dissociation constant of 19 μM (Figure 2). These values are comparable to those of natural TPRs binding their cognate peptide ligands. For example, wild-type TPR2A binds the C-terminal peptide of Hsp90 with a dissociation constant of 6 μM and TPR1 (also from Hop) binds the C-terminal peptide of Hsp70 with a dissociation constant of 18 μM (*15*).

For the proposed *in vivo* applications, it was important to confirm that the T-Mods recognize not only the Dss1-C19 peptide but also the full-length Dss1 protein. We therefore tested the ability of the T-Mods to bind to full-length human Dss1, fused to GST for ease of purification and identification. Both T-Mod(Dss1A) and T-Mod(Dss1B) bind to full-length Dss1 in this context and show no nonspecific binding to GST alone or to GST fused to a noncognate peptide (data not shown).

To obtain a better molecular understanding of the T-Mod-Dss1-C19 interaction, we sought to map the region within the Dss1-C19 peptide recognized by the T-Mods. The paradigm for TPR-peptide recognition is that 5−7 residues of the peptide are bound in the TPR recognition pocket in an extended conformation. With this in mind, we synthesized three shorter peptides: two that were the N- and C-terminal halves of the 19-mer Dss1-C19 peptide and a third that covered the central region. We assayed binding to the T-Mods by ITC and observed no binding (data not shown). We then explored the possibility that the Dss1-C19 peptide may have secondary structure critical for the interaction. The circular dichroism spectra of the Dss1-C19 peptide, both free and bound to the T-Mod, indicate that it is helical (Figure 3, panel a). We also used the program Agadir to predict the helical content of the Dss1 peptide (http://www.embl-heidelberg.de/Services/serrano/agadir/agadir-start.html) (Figure 4, panel a) (*40*). The Agadir predictions are in agreement not only with the percent helicity of the Dss1-C19 peptide but also with the region of the Dss1 sequence that is helical in the co-crystal structure with BRCA2 (Figure 4, panel a and Figure 2, panel b).

We generated a working structural model for the Dss1-TPR interaction, which integrates the information on both the Dss1-C19 peptide and the TPR binding surface. First, we know that Dss1-C19 adopts a helical structure, both in solution and in complex with the T-Mods. Second, we know the region of Dss1 that is predicted to be helical in the free peptide and that this region is also helical in the BRCA2-Dss1

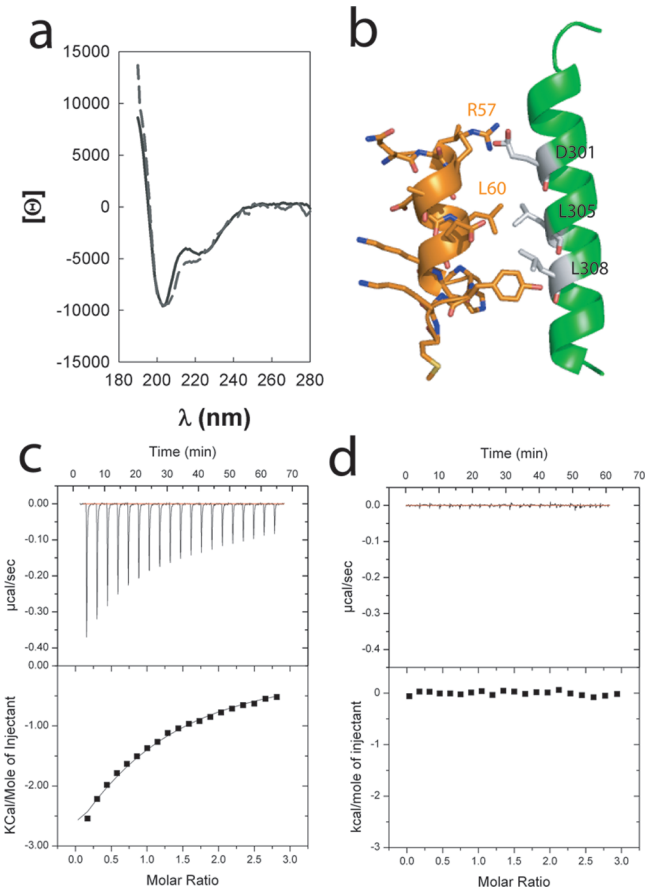

Figure 3. T-Mods-Dss1 binding mode. a) Circular dichroism spectra of Dss1-C19 peptide free in solution at 50 μM (solid line) and bound to T-Mod(Dss1A) (dashed line). The spectrum of the bound peptide corresponds to the subtraction of the T-Mod(Dss1A) spectrum (50 μM) from the T-Mod(Dss1A) in complex with Dss1-C19 peptide (50 μM each). b) Structural model of the T-Mod(Dss1)-Dss1 peptide binding interface. The first helix of the third repeat of the TPR module is shown as a green ribbon. The three randomized residues that reached a consensus in the Dss1-binding modules are shown as gray sticks. The C-terminal part of Dss1 from the Dss1-BRCA2 complex structure is shown in orange. The amino acid side chains are shown in sticks. The key residues for the Dss1-C19-TPR interaction, R57 and L60, are labeled. c) ITC binding isotherms for the interaction of Dss1-C19 peptide mutant R57A with T-Mod(Dss1A). The Dss1-C19 peptide was titrated at 1.37 mM concentration into 100 μM T-Mod(Dss1A) solution. The binding enthalpies were integrated and fit to a 1:1 binding model to calculate the stoichiometry (*N*) and the binding affinity (K_d) of the interactions (T-Mod(Dss1A) K_d = 157 μM, N = 0.8). d) ITC binding isotherms for the interaction of Dss1-C19 peptide mutant L60A with T-Mod(Dss1A). The Dss1-C19 peptide was titrated at 1.32 mM into 100 μM T-Mod(Dss1A) solution.

co-crystal structure. Third, we know the consensus residues for Dss1 binding on the T-Mod. Thus, it is reasonable to model Dss1 as binding to the T-Mod in a helical conformation, with the conserved residues identified in the screen playing a key role in Dss1 binding. The working model shown in Figure 3, panel b represents the most reasonable solution using these constraints. We generated a series of Dss1-C19 mutants to test this model. We individually mutated the residues R57 and L60, which were predicted to be critical for the Dss1-T-Mods interaction, to Ala, so as not to perturb the helicity of the peptide (which was confirmed by CD; data not shown). We measured the binding to the T-Mods(Dss1A and B) by ITC. The effects of the mutations on the Dss1-C19 peptide were comparable for both T-Mods. The R57A mutation significantly reduced the binding affinity (Figure 3, panel c) ($K_{d\ T\text{-}Mod(Dss1A)}$ = 157 μM; $K_{d\ T\text{-}Mod(Dss1B)}$ = 665 μM), and the L60A mutation, which is positioned in the central packing layer of the predicted interaction surface, completely disrupted the binding for both T-mods (Figure 3, panel d). The results of the point mutations support the working model and emphasize the specificity of the T-Mod−Dss1 interaction.

Using T-Mods To Perturb Cellular Pathways in Yeast. Yeast Sem1 has 40% overall sequence identity and 70% sequence similarity with the human Dss1 protein (ClustalW (*41*)) (Figure 4, panel a). We used ITC to assess the ability of the anti-Dss1 T-Mods to bind to Sem1 *in vitro*. ITC experiments confirmed the binding of the T-Mods to Sem1, albeit with weaker affinity than Dss1 (K_d = 175 μM) (data not shown).

The biological functions of *Saccharomyces cerevisiae* Sem1 have been studied extensively (*26, 32–34*). A strain of *S. cerevisiae* lacking the gene encoding Sem1 (*sem1*Δ) exhibits no significant growth defect at 25 °C, but is nonviable at 37 °C (*26*). The *sem1*Δ strain is complemented equally well by plasmid expressed yeast Sem1 or human Dss1 (Figure 4, panel b) (*34*). With these observations as background, we tested the ability of the T-Mods to bind and thereby inhibit Dss1 or Sem1 function *in vivo*. When the anti-Dss1 T-Mods are expressed in yeast, there is no observable growth defect at 25 °C, *i.e.*, expression

of the T-Mods has no general detrimental effect (Figure 4, panel b). We have expressed several other T-Mods with different binding specificities in yeast and observed no growth defect whatsoever at either 25 or 37 °C. At 37 °C, expression of the anti-Dss1 T-Mod results in cell death (Figure 4, panel b). We interpret this result as being consistent with binding to and thereby inhibition of endogenous Sem1 activity. Further support for this interpretation comes from the finding that supplying exogenous Dss1 in cells expressing the T-Mod partially reverses the growth defect at 37 °C; this presumably reflects the partial titration of the dominant-negative T-Mod through binding to Dss1. The severity of the growth defect directly correlates with the expression level of the T-Mods: a more severe defect is observed when the T-Mods are expressed from a high-copy number plasmid (p424-GAL1) compared to when they are expressed from a low-copy plasmid (p414-GAL1, data not shown). This effect is observed either for Wt or for *sem1*Δ complemented with Sem1 or Dss1, the effects of the T-Mods being less severe when *sem1*Δ is complemented by plasmid, presumably because of the higher expression levels of Sem1 and Dss1.

A temperature-sensitive growth defect does not, alone, prove that the T-Mods are interfering with normal Sem1 function *in vivo* as opposed to having a nonspecific toxic effect at elevated temperature. We therefore devised a genetic test that specifically examines whether the T-Mod impairs proteasome-associated Sem1 activity. Previous work has shown that cells lacking both the proteasomal subunit Rpn10 and Sem1 (*rpn10*Δ*sem1*Δ) are extremely sensitive to any stress that requires proteasome activity. One such stress is DNA damage induced by hydroxyurea (HU). *rpn10*Δ*sem1*Δ is killed in the presence of HU, whereas the *rpn10*Δ strain is not. Sem1 and the Rpn10 proteasome subunit are thus both required to maintain functionality of the proteasome (*33*). When the anti-Dss1 T-Mod is expressed in the *rpn10*Δ strain (Figure 4, panel c) and the cells are exposed to DNA-damaging concentrations of hydroxyurea (HU), the cells die. Thus, expression of the anti-Dss1 TPR, in the context of the *rpn10*Δ, recapitulates the effect of a chromosomal deletion of Sem1. The effect is specific: there is no sensitivity to HU at 30 °C if the anti-Dss1 TPR is expressed in a wild-type strain or if the T-Mod is expressed in *rpn10*Δ cells at 30 °C, in the absence of HU. The sensitivity to HU in *rpn10*Δ expressing anti-Dss1 TPR is marked, but less severe than the sensitivity seen with *rpn10*Δ*sem1*Δ cells. This observation is consistent with the T-Mod inhibiting Sem1 proteasome associated function. The ability of the anti-Dss1 T-Mod to phenocopy a specific *sem1*Δ *rpn10*Δ defect directly links the expression of T-Mods to inhibition of the proteasomal function. In particular, the anti-Dss1 T-Mods have a pronounced effect on the proteasome-associated function in DNA repair (*42*).

We have described a generally applicable method by which to create novel binding modules to study protein function *in vivo*. Such modules have a variety of applications, such as rewiring cellular networks or tracking or inhibiting cellular proteins, which we have demonstrated. An important feature of such genetically encoded modules is that their intracellular concentrations can be finely tuned, both temporally and spatially. By contrast, the use of exogenously added small molecules is limited by the uncertainly of intracellular concentrations, which depend both on the cellular uptake and metabolic degradation of the compound. The T-Mod strategy is also advantageous over methods, such as siRNA, that simply lower or abolish the expression of a particular protein because it allows the inhibition of specific interactions of a protein, while leaving other activities intact.

METHODS

TPR Library Construction and Library Screen. Seven residues were selected for randomization (K229, N233, Y236, N264, K301, R305, and N308 of Hsp Organizing protein, Hop, where the TPR2A domain is amino acids 222−352). To construct the library, we designed six overlapping oligonucleotides corresponding to the TPR2A gene and assembled a combinatorial library of genes by Klenow extension and PCR (*43*). We randomized the selected positions by using a codon mixture (in the oligo synthesis) with an equal, unbiased distribution of 19 amino acids (cysteine codons were excluded from the mixture to avoid the formation of disulfide linkages). The gene library was cloned into the pMRBAD-CGFP vector expressing the C-terminal half of GFP and transformed in *E. coli* cells containing the DSS1-C19 target peptide fused to the N-terminal half of GFP. After growth and induction the cell population was screened using FACS, and the fluorescent-positive colonies were isolated and sequenced. We provide a full description of the experimental conditions in the following Article.

TPR Cloning, Purification, and Characterization. Two of the positive hits identified in the screen were re-cloned into the pPro-Ex-HTa vector for protein overexpression. The proteins were expressed as His-tagged fusions and purified using standard protocols. The circular dichroism spectra and the thermal denaturation curves of the proteins were acquired as described in the following Article.

Isothermal Titration Calorimetry (ITC) Binding Assays. Experiments were performed using a Microcal ITC-200 in 150 mM NaCl, 50 mM phosphate, pH 7.4, at 25 °C. The reference power was set to 4 μCal s^{-1}, and injection volume to 2 μL with 180 s equilibration time between injections. Dss1-C19 peptide solutions were titrated at a concentration from 1 to 2 mM into the T-Mod(Dss1) solutions at a protein concentration around 100 μM. The binding enthalpies were integrated and fit to a 1:1 binding model to calculate the stoichiometry (*N*) and the binding affinity (K_d) of the interactions using Origin 7.0.

Yeast Strains, Culture Conditions, and Genetic Methods. *S. cerevisiae* YPH499 is the WT strain, YMF12-1a (MHY2835) is the strain lacking the chromosomal SEM1 gene (*sem1*Δ), MHY4550 is the strain lacking the chromosomal RPN10 gene (*rpn10*Δ), and MHY-5817 is the congenic strain lacking both SEM1 and RPN10 genes (*sem1*Δ*rpn10*Δ) (*33*). Full-length dss1 protein was cloned into the p416GPD vector (URA3, GPD promotor). TPR domains T-Mod(Dss1A) and T-Mod(Dss1B) were cloned fused to a N-terminal nuclear localization signal (SV40-NLS) in p414-GAL1 (TRP1, CEN expression vector, GAL1 promoter) and p424-GAL1 (TRP1, 2 μm expression vector, GAL1 promoter) vectors. Synthetic media (SD) lacking

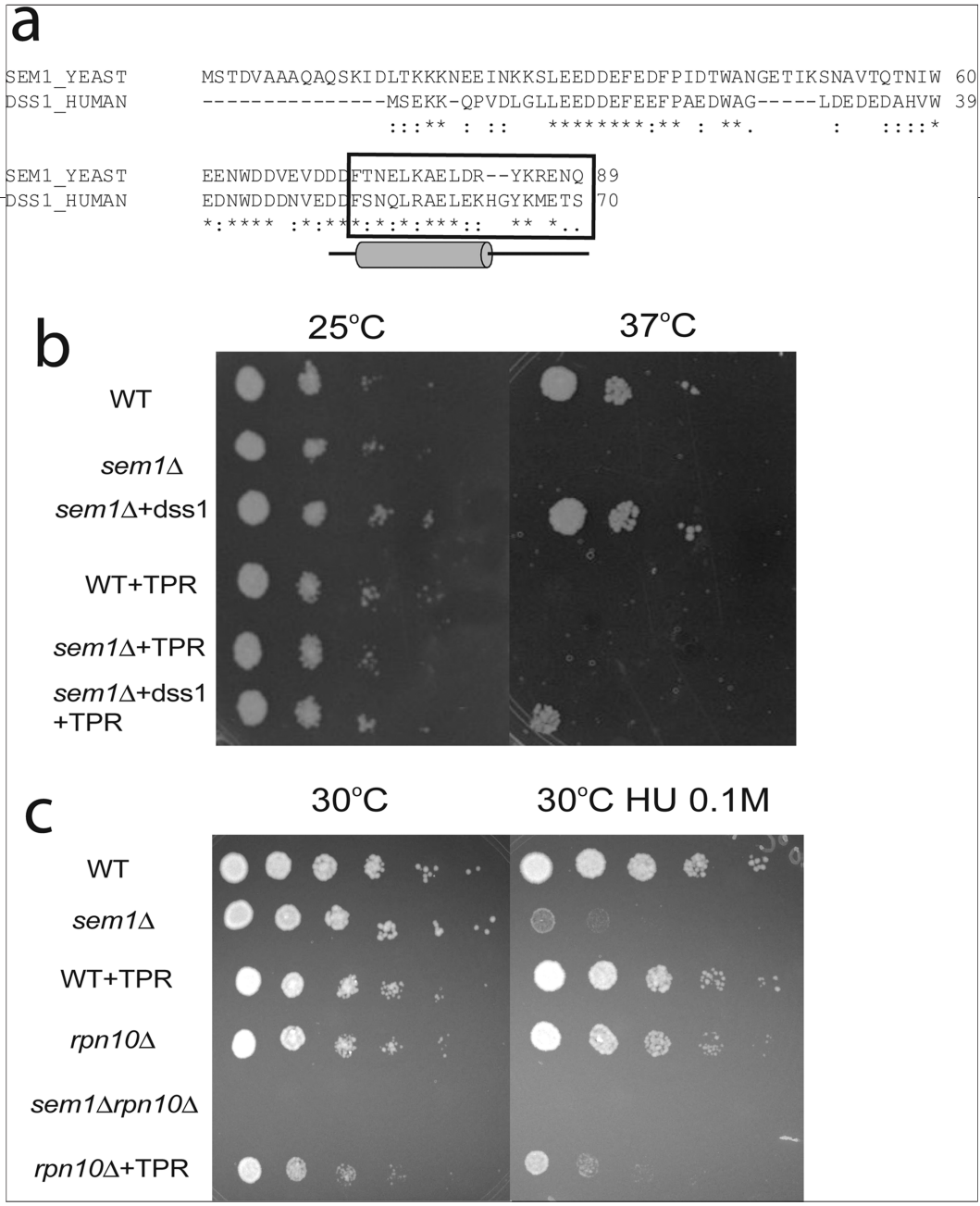

Figure 4. ***In vivo*** **binding activity of TPR domains. a) Sequence alignment of human Dss1 protein and the yeast ortholog Sem1, where "*" means identical residues, and ":" means similar residues. The black square highlights the 19 C-terminal residues of Dss1 protein that were targeted during the selection. The schematic shows the results of the secondary structure prediction in the C-terminal region by Agadir (*40*). b) Temperature-sensitive growth phenotype of *S. cerevisiae* transformants. Cell cultures were diluted to the same concentration, and serially diluted cells were spotted onto galactose plates and incubated at 25 or 37 °C. The strains of *S. cerevisiae*, plated from top to bottom, are the following: wild-type (WT), *sem1*Δ (cells lacking the chromosomal *SEM1* gene), *sem1*Δ transformed with the plasmid pGPD416-Dss1 (*sem1*Δ+Dss1), WT expressing T-Mod(Dss1A) fused to a N-terminal nuclear localization signal (SV40-NLS) from a p424-GAL1 vector (WT+TPR), *sem1*Δ expressing T-Mod(Dss1A) fused to SV40-NLS from the p424-GAL1 vector (*sem1*Δ+TPR), *sem1*Δ complemented with pGPD416-Dss1 and expressing T-Mod(Dss1A) with a N-terminal NLS from the p424-GAL1 vector (*sem1*Δ +Dss1+TPR). c) Hydroxyurea-sensitive growth phenotype of *S. cerevisiae* transformants. Cell cultures were diluted to the same concentration, and serially diluted cells were spotted into SD galactose plates and SD galactose plates with 0.1 M hydroxyurea (HU) and incubated at 30 °C. The strains of *S. cerevisiae*, plated from top to bottom, are the following: wild-type (WT), *sem1*Δ (cells lacking the chromosomal *SEM1* gene), WT expressing T-Mod(Dss1A) fused to a N-terminal nuclear localization signal (SV40-NLS) from a p424-GAL1 vector (WT+TPR), *rpn10*Δ (cells lacking the chromosomal *RPN10* gene), *sem1*Δ *rpn10*Δ (cells lacking both the chromosomal *SEM1* and *RPN10* genes), *rpn10*Δ expressing T-Mod(Dss1A) fused to SV40-NLS from the p424-GAL1 vector (*rpn10*Δ+TPR).**

appropriate nutrient(s) were used to select strains containing specific plasmids.

For the growth assays the strains were cultured overnight in SD media lacking the appropriate nutrients at 25 °C. The cell density was adjusted to A_{600} = 0.5, and 1.5 μL of 5-fold serially diluted cells was spotted onto galactose plates and incubated at 25 or 37 °C. For the hydroxyurea-sensitive growth assay, cell cultures were diluted to the same concentration and serially diluted cells were spotted onto SD galactose plates and SD galactose plates with 0.1 M hydroxyurea (HU) and incubated at 30 °C.

Acknowledgment: We thank M. Funakoshi for his valuable advice with yeast experiments. We thank G. Lyon at FACS Facility at Yale University School of Medicine for assistance with FACS experiment. We thank P. Sung Lab (Yale University) for providing the GST-Dss1 plasmid and purification protocol and anti-Dss1 antibodies. We thank R. Collins, C. Cheng, T. Z. Grove, R. Ilagan, M. E. Jackrel, L. Kundrat, and G. Pimienta-Rosales for valuable discussions and comments on the manuscript. T.Y.L. was a recipient of Yale-HHMI Future Scientist Summer Fellowship, Michael Manzella Foundation Summer Fellowship, Sheffield Society Undergraduate Research Fellowship, and Science, Technology, and Research Scholars (STARS) II Fellowship. Financial support for M.H. from National Institutes of Health grant number R01 GM083050 is acknowledged. This research is funded in part by NIH and HFSP.

REFERENCES

1. Consortium, I. H. G. S. (2004) Finishing the euchromatic sequence of the human genome, *Nature 431*, 931–945.

2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., and Yan, C., *et al.* (2001) The sequence of the human genome, *Science 291*, 1304–1351.
3. Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrin-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. J., Musso, G., St Onge, P., Ghanny, S., Lam, M. H., Butland, G., Altaf-Ul, A. M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*, *Nature 440*, 637–643.
4. Walhout, A. J., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., and Vidal, M. (2000) Protein interaction mapping in *C. elegans* using proteins involved in vulval development, *Science 287*, 116–122.
5. Barrios-Rodiles, M., Brown, K. R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R. S., Shinjo, F., Liu, Y., Dembowy, J., Taylor, I. W., Luga, V., Przulj, N., Robinson, M., Suzuki, H., Hayashizaki, Y., Jurisica, I., and Wrana, J. L. (2005) High-throughput mapping of a dynamic signaling network in mammalian cells, *Science 307*, 1621–1625.
6. Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P., and Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network, *Nature 437*, 1173–1178.
7. Mirecka, E. A., Hey, T., Fiedler, U., Rudolph, R., and Hatzfeld, M. (2009) Affilin molecules selected against the human papillomavirus E7 protein inhibit the proliferation of target cells, *J. Mol. Biol. 390*, 710–721.
8. Bashor, C. J., Helman, N. C., Yan, S., and Lim, W. A. (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics, *Science 319*, 1539–1543.
9. Gebauer, M., and Skerra, A. (2009) Engineered protein scaffolds as next-generation antibody therapeutics, *Curr. Opin. Chem. Biol. 13*, 245–255.
10. D'Andrea, L., and Regan, L. (2003) TPR proteins: the versatile helix, *Trends Biochem. Sci. 28*, 655–662.
11. Lamb, J. R., Tugendreich, S., and Hieter, P. (1995) Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem. Sci. 20*, 257–259.
12. Das, A. K., Cohen, P. W., and Barford, D. (1998) The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions, *EMBO J. 17*, 1192–1199.
13. Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U., and Moarefi, I. (2000) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine, *Cell 101*, 199–210.
14. Gatto, G. J., Jr., Geisbrecht, B. V., Gould, S. J., and Berg, J. M. (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5, *Nat. Struct. Biol. 7*, 1091–1095.
15. Brinker, A., Scheufler, C., Von Der Mulbe, F., Fleckenstein, B., Herrmann, C., Jung, G., Moarefi, I., and Hartl, F. U. (2002) Ligand discrimination by TPR domains. Relevance and selectivity of EEVD-recognition in Hsp70 x Hop x Hsp90 complexes, *J. Biol. Chem. 277*, 19265–19275.
16. Main, E. R. G., Xiong, Y., Cocco, M. J., D'Andrea, L., and Regan, L. (2003) Design of stable α-helical arrays from an idealized TPR motif, *Structure 11*, 497–508.
17. Main, E. R. G., Jackson, S. E., and Regan, L. (2003) The folding and design of repeat proteins: reaching a consensus, *Curr. Opin. Struct. Biol. 13*, 482–489.
18. Kajander, T., Cortajarena, A. L., Mochrie, S. G., and Regan, L. (2007) Structure and stability of a consensus TPR superhelix, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 63*, 800–811.
19. Cortajarena, A. L., and Regan, L. (2006) Ligand binding by TPR domains, *Protein Sci. 15*, 1193–1198.
20. Cortajarena, A. L., Yi, F., and Regan, L. (2008) Designed TPR modules as novel anticancer agents, *ACS Chem. Biol. 3*, 161–166.
21. Cortajarena, A. L., Kajander, T., Pan, W., Cocco, M. J., and Regan, L. (2004) Protein design to understand peptide ligand recognition by tetratricopeptide repeat proteins, *PEDS 17*, 399–409.
22. Jackrel, M. E., Valverde, R., and Regan, L. (2009) Redesign of a protein-peptide interaction: characterization and applications, *Protein Sci. 18*, 762–774.
23. Cortajarena, A. L., Wang, J., and Regan, L. (2009) Crystal structure of a designed TPR module in complex with its peptide-ligand, *FEBS J. 277*, 1058–1066.
24. Kajander, T., Sachs, J. N., Goldman, A., and Regan, L. (2009) Electrostatic interactions of Hsp-organizing protein tetratricopeptide domains with Hsp70 and Hsp90: computational analysis and protein engineering, *J. Biol. Chem. 284*, 25364–25374.
25. Crackower, M. A., Scherer, S. W., Rommens, J. M., Hui, C. C., Poorkaj, P., Soder, S., Cobben, J. M., Hudgins, L., Evans, J. P., and Tsui, L. C. (1996) Characterization of the split hand/split foot malformation locus SHFM1 at 7q21.3-q22.1 and analysis of a candidate gene for its expression during limb development, *Hum. Mol. Genet. 5*, 571–579.
26. Marston, N. J., Richards, W. J., Hughes, D., Bertwistle, D., Marshall, C. J., and Ashworth, A. (1999) Interaction between the product of the breast cancer susceptibility gene BRCA2 and DSS1, a protein functionally conserved from yeast to mammals, *Mol. Cell. Biol. 19*, 4633–4642.
27. Sung, P., and Klein, H. (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions, *Nat. Rev. Mol. Cell Biol. 7*, 739–750.
28. Li, J., Zou, C., Bai, Y., Wazer, D. E., Band, V., and Gao, Q. (2006) DSS1 is required for the stability of BRCA2, *Oncogene 25*, 1186–1194.
29. Kojic, M., Yang, H., Kostrub, C. F., Pavletich, N. P., and Holloman, W. K. (2003) The BRCA2-interacting protein DSS1 is vital for DNA repair, recombination, and genome stability in Ustilago maydis, *Mol. Cell 12*, 1043–1049.
30. Yang, H., Jeffrey, P. D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N. H., Zheng, N., Chen, P. L., Lee, W. H., and Pavletich, N. P. (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure, *Science 297*, 1837–1848.
31. Thakurta, A. G., Gopal, G., Yoon, J. H., Kozak, L., and Dhar, R. (2005) Homolog of BRCA2-interacting Dss1p and Uap56p link Mlo3p and Rae1p for mRNA export in fission yeast, *EMBO J. 24*, 2512–2523.
32. Wilmes, G. M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S. R., Whitworth, G. B., Kress, T. L., Weissman, J. S., Ideker, T., Guthrie, C., and Krogan, N. J. (2008) A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing, *Mol. Cell 32*, 735–746.
33. Funakoshi, M., Li, X., Velichutina, I., Hochstrasser, M., and Kobayashi, H. (2004) Sem1, the yeast ortholog of a human BRCA2-binding protein, is a component of the proteasome regulatory particle that enhances proteasome stability, *J. Cell Sci. 117*, 6447–6454.
34. Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004) Sem1p is a novel subunit of the 26 S proteasome from *Saccharomyces cerevisiae*, *J. Biol. Chem. 279*, 28807–28816.
35. Josse, L., Harley, M. E., Pires, I. M., and Hughes, D. A. (2006) Fission yeast Dss1 associates with the proteasome and is required for efficient ubiquitin-dependent proteolysis, *Biochem. J. 393*, 303–309.
36. Magliery, T. J., and Regan, L. (2005) Sequence variation in ligand binding sites in proteins, *BMC Bioinf. 6*, 240.

37. Magliery, T. J., and Regan, L. (2004) Beyond consensus: statistical free energies reveal hidden interactions in the design of a TPR motif, *J. Mol. Biol. 343*, 731–745.
38. Magliery, T. J., Wilson, C. G. M., Pan, W. L., Mishler, D., Ghosh, I., Hamilton, A. D., and Regan, L. (2005) Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism, *J. Am. Chem. Soc. 127*, 146–157.
39. Wilson, C., Magliery, T., and Regan, L. (2004) Detecting protein-protein interactions with GFP-fragment reassembly, *Nat. Methods 1*, 255–262.
40. Munoz, V., and Serrano, L. (1997) Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms, *Biopolymers 41*, 495–509.
41. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and D.G., H. (2007) ClustalW and ClustalX version 2, *Bioinformatics 23*, 2947–2948.
42. Krogan, N. J., Lam, M. H., Fillingham, J., Keogh, M. C., Gebbia, M., Li, J., Datta, N., Cagney, G., Buratowski, S., Emili, A., and Greenblatt, J. F. (2004) Proteasome involvement in the repair of DNA double-strand breaks, *Mol. Cell 16*, 1027–3104.
43. Kajander, T., Cortajarena, A. L., and Regan, L. (2006) Consensus design as a tool for engineering repeat proteins, *Methods Mol. Biol. 340*, 151–170.