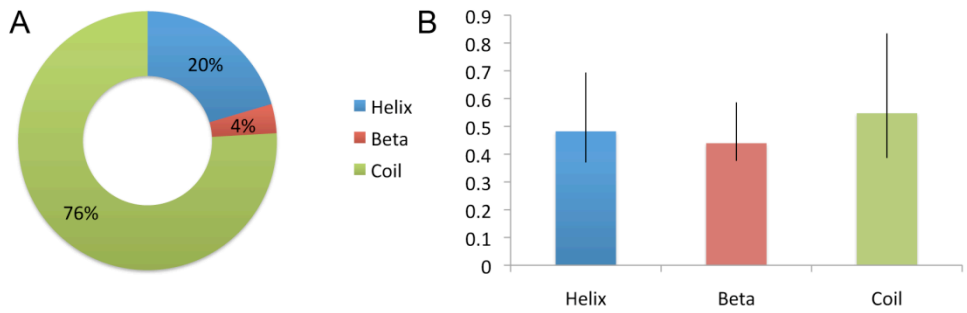

Figure S1: Secondary structure of high-affinity peptides derived from Benchmark 3.0 complexes: β-strands are underrepresented. A peptide was classified to a secondary structure category, if more than 4 of its residues adopt the given secondary structure (secondary structures were defined according to Stride[1]). **A.** Distribution of secondary structure categories in the derived high-affinity peptides. Peptides complementing a β-sheet are under-represented in comparison to natively solved peptide-protein complexes[2]. **B.** Relative contribution to total binding energy for peptides of different secondary structure types (median values; error bars represent percentiles of 15% and 85%, respectively). There is no significant difference between values for the different categories of secondary structure.

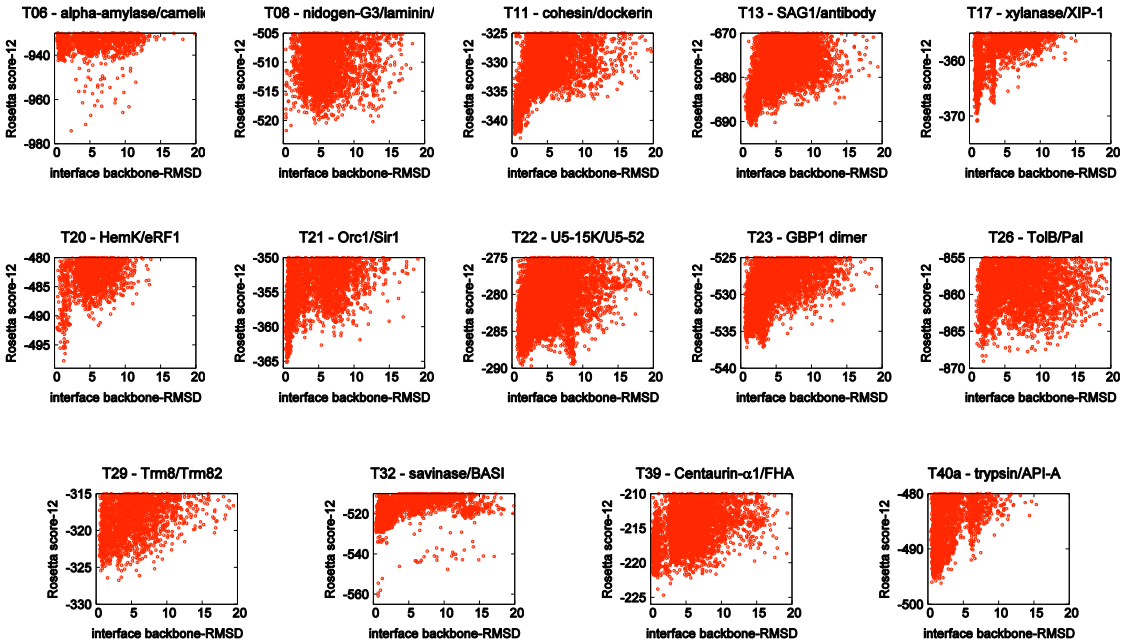

Figure S2: Energy funnels for peptides derived from 14 CAPRI targets indicate that peptides strongly favor their original binding conformation in their origin protein. For each interaction, 9000 models were generated by extensive flexible peptide docking (see Methods). Each docking model is presented as a dot, indicating its total energy (*y*-axis) *vs.* the backbone RMSD of peptide interface residues in the native protein context (*x*-axis). For clarity, only top-scoring models out of the total 9000 are shown.

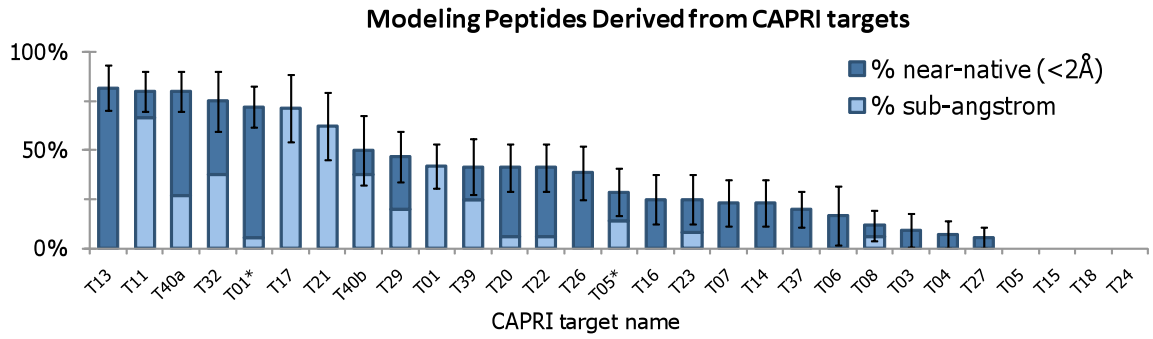

Figure S3: Derived peptides sample the native conformation observed in the context of the original complex. For peptides derived from each CAPRI target, a wide range of perturbed starting structures were created by adding strong random noise to the native peptide conformation, followed by extensive peptide docking (see Methods). The plot shows the percent of successful docks for structures with initial perturbations of 2Å-5Å backbone-RMSD from the native conformation of the derived peptides, which is within the basin of attraction of the FlexPepDock modeling protocol[3]. The bars indicate successful docks in which the top-ranking decoy had sub-Angstrom modeling accuracy (<1Å interface bb-RMSD to the native conformation of the derived peptide; light-blue) or near-native modeling accuracy (<2Å interface bb-RMSD). For most CAPRI targets, the derived peptide favors its cognate binding mode in the context of the original complex.

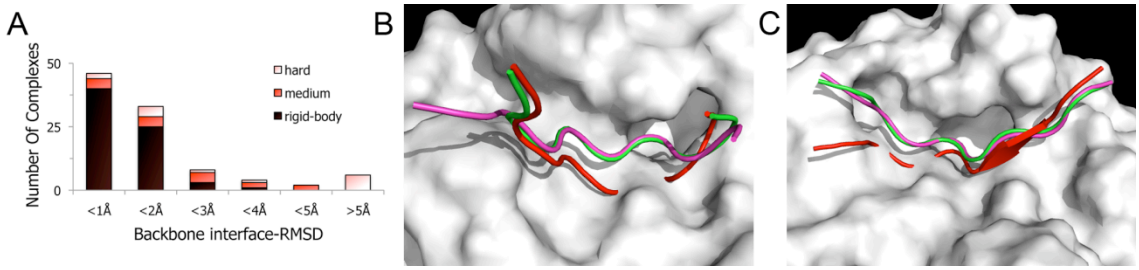

Figure S4: Native binding conformations of derived peptides can be recovered from the conformation of the peptide in the unbound monomer by peptide docking. A. Plot of the distribution of backbone RMSD values over the derived peptide residues between the bound and free conformation of the original globular domain from which the peptide was derived. The different Benchmark 3.0 difficulty classes for the original interaction (as defined by Hwang *et al.* [4]; see legend) are plotted separately. **B-C.** Examples of successful docking simulations on peptides derived from the following complexes: **B.** Trypsin/Tryptase inhibitor from tick (pdb: 2UUY). **C.** Chymotrypsin/Ecotin (pdb: 1N8O). The native peptide is colored green, the peptide in its unbound form (as part of the unbound monomer) is colored red, and the model created with the FlexPepDock protocol is colored magenta.

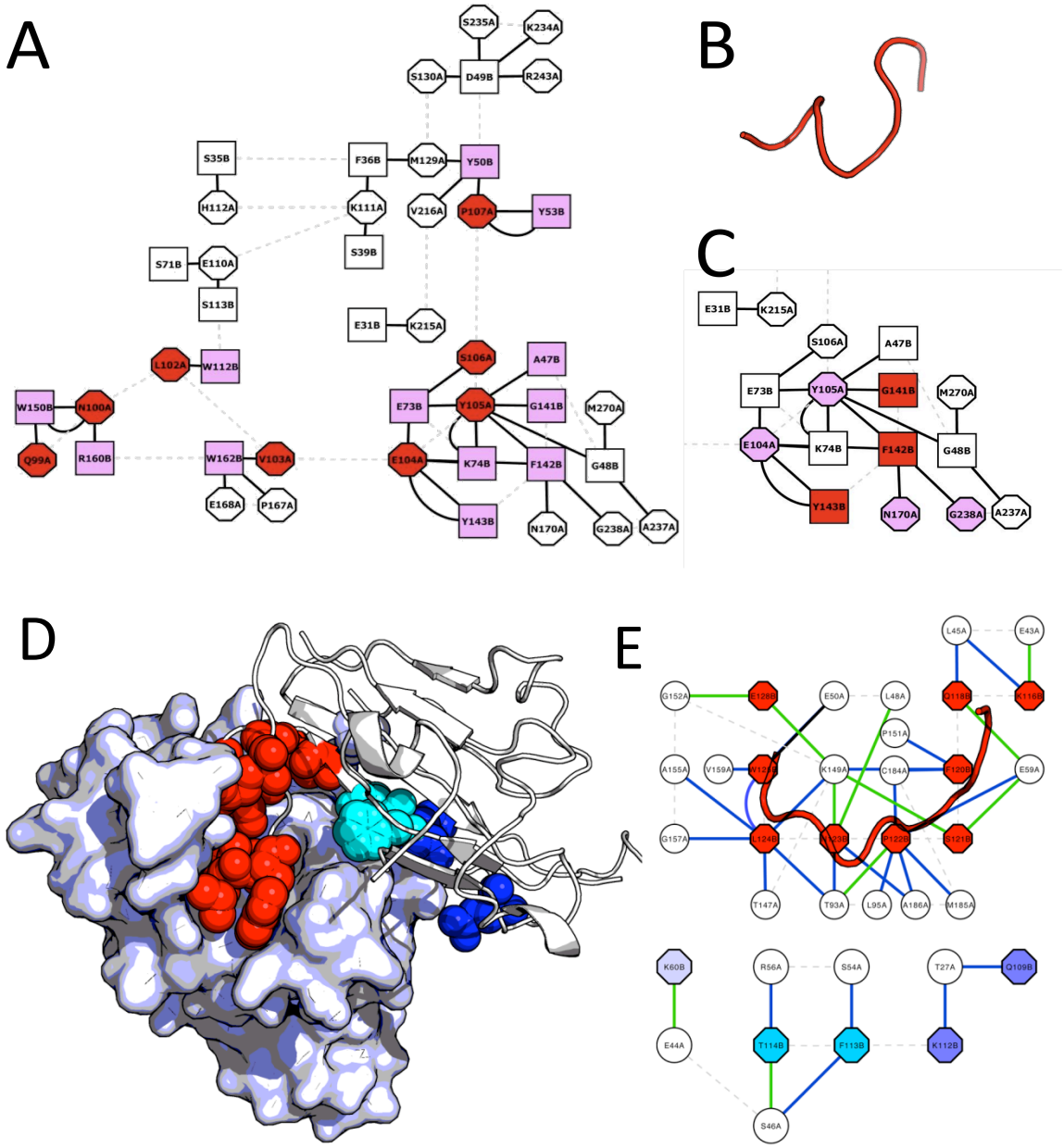

Figure S5. Analysis of contribution of derived high-affinity peptides to different binding modules at the interface. In an analysis of the network of interactions at the interface of the interaction between TEM1-β-lactamase and its inhibitor β-lactamase inhibitor protein BLIP, Reichman *et al.*[5,6] demonstrated that protein interfaces may consist of several independent binding modules. Their experimental assessment indicated that a mutation within one module will not affect the other modules, while mutations within the same module tend to be non-additive. We assess here the contribution of a high-affinity derived peptide to this nodes in two cases: TEM1-BLIP **(A-C.)** and EphB4-EphrinB2 **(D–E).** Networks were visualized using the Cytoscape software[7]. The linear peptide is marked by red nodes in all schemes. Covalent bonds

are depicted as gray dashed lines (including residues that are two amino acids apart). Non-bonded interactions were identified using the AquaProt server, as described in Reichman *et al.*[6], and are depicted as solid lines.

A-C. High-affinity peptides in the interface of TEM1-BLIP contribute to one or more binding modules. **A.** The BLIP side of the interface: A schematic view of the network of interactions between residues of TEM1 (octagons) and β-Lactamase Inhibitor Protein (BLIP; squares) indicates that the highest-affinity linear segment on the interface of BLIP (residues 99-108; red) contributes to several different binding modules, and seems to serve as a "hub of modules". **B.** The sharp turns in the coiled structure of the highest-affinity peptide of BLIP (residues 99-108) may explain its contribution to independent binding modules. **C.** The TEM1 side of the interface: A three-residue stretch on the interface of TEM1 (residues 142-144; red - note the change in the color code compared to A.) constitutes a binding "hot-segment" in the biggest binding module of the TEM1-BLIP network. These three residues are responsible for 41% of the interface total binding energy, although they participate in only one of the five binding modules. **D-E. The proposed inhibitory peptide derived from EphrinB2 constitutes the largest interface cluster in the network of the EphB4-EphrinB2 interaction. D.** Structural representation: EphB4 is depicted in surface representation and EphrinB2 as cartoon, and different interface clusters (donated by EphrinB2) are colored by cluster identity and represented in spacefill (the red cluster covers the suggested high-affinity linear peptide at the G-H loop, see text). **E.** Schematic presentation of the network of interactions between residues of EphB4 (circles) and EphrinB2 (octahedrons). EphrinB2 residues are colored by their interface clusters with corresponding colors. Here, non-covalent interactions are colored as: blue – vdW interactions; green – hydrogen-bonds; black – aromatic interactions.

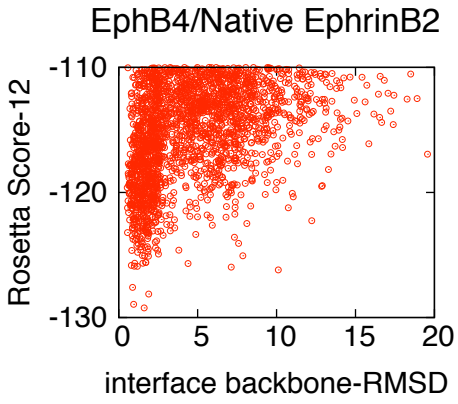

Figure S6: The energy landscape of the EphrinB2 derived peptide. A pronounced energy funnel is found leading towards the native binding mode. Note that this funnel is for the lead peptide, prior to the subsequent design steps.

References:

1. Frishman D, Argos P. Knowledge-based protein secondary structure assignment. Proteins 1995;23(4):566-579.
2. London N, Movshovitz-Attias D, Schueler-Furman O. The structural basis of peptide-protein binding strategies. Structure 2010;18(2):188-199.
3. Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins: Structure, Function, and Bioinformatics 2010;78:2029-2040.
4. Hwang H, Pierce B, Mintseris J, Janin J, Weng Z. Protein-protein docking benchmark version 3.0. Proteins 2008;73:705-709.
5. Reichmann D, Rahat O, Albeck S, Meged R, Dym O, Schreiber G. The modular architecture of protein-protein binding interfaces. Proceedings of the National Academy of Sciences of the United States of America 2005;102:57-62.
6. Reichmann D, Cohen M, Abramovich R, Dym O, Lim D, Strynadka NCJ, Schreiber G. Binding hot spots in the TEM1-BLIP interface in light of its modular architecture. Journal of Molecular Biology 2007;365:663-679.
7. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498-2504.