

RESEARCH ARTICLE SUMMARY

PROTEIN INTERACTIONS

Structural ontogeny of protein-protein interactions

Aerin Yang†, Hanlun Jiang†, Kevin M. Jude†, Deniz Akpinaroglu, Stephan Allenspach†, Alex Jie Li, James Bowden, Carla Patricia Perez, Liu Liu, Po-Ssu Huang, Tanja Kortemme, Jennifer Listgarten*, K. Christopher Garcia*

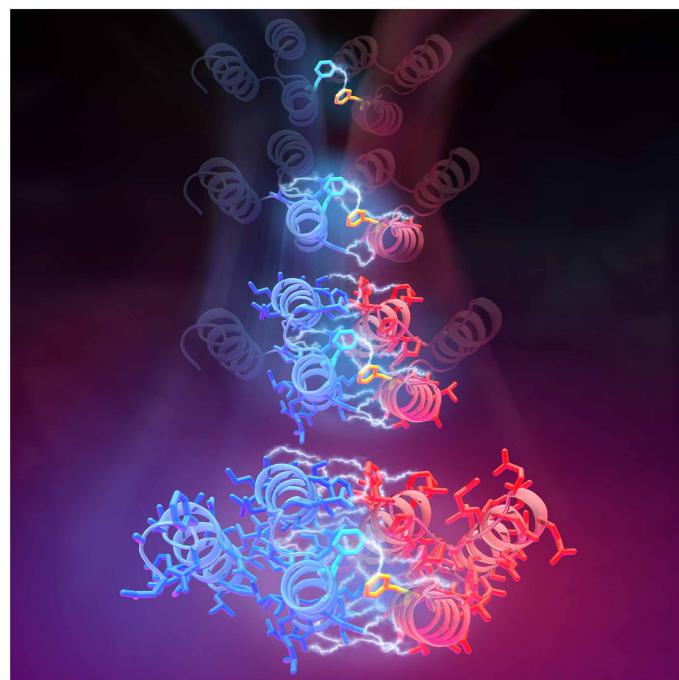


Full article and list of author affiliations:
<https://doi.org/10.1126/science.adx6931>

INTRODUCTION: How proteins initiate and evolve interactions with other proteins to form binding sites and protein-protein interfaces remains poorly understood. Yet there appear to be distinctive properties of protein surfaces that have evolved to form protein-protein interfaces versus noninteracting surfaces of proteins. Natural protein binding sites are usually the most “druggable” sites on a given protein and tend to attract the majority of binders from combinatorial peptide or antibody libraries. Are there biophysical properties that distinguish regions of protein surfaces that have not evolved to bind to a ligand from natural binding sites? This question has implications for drug discovery.

RATIONALE: In this study, we asked whether there are fundamental differences between surfaces of proteins that have evolved to bind to proteins versus surfaces that have not. To do this, we carried out a synthetic proxy of protein-protein coevolution by designing completely new interfaces between “silent” surfaces of proteins that have no ligand recognition history and do not normally interact. We experimentally simulated and then computationally deconstructed the process by which protein-protein interactions are initially formed and subsequently coevolve specificity and high affinity. This synthetic proxy of protein coevolution has also allowed us to gain a better insight into the differences between natural and nonnatural binding sites of proteins.

RESULTS: By using a synthetic coevolution platform to design interactions between naïve surfaces unencumbered by a natural binding imprint, we captured previously unseen snapshots illustrating the formation of de novo protein complexes. We found that the type of epistasis and the structural adaptations that shape affinity, specificity, and orthogonality differ between complexes formed by surfaces that have naturally evolved to bind other proteins versus nonbinding surfaces. Remodeling an interface formed by natural surfaces led to high-affinity solutions with distinct specificities, but the docking mode stayed the same in a deep energy well. Selecting for completely new interfaces mediated by a “naïve” surface of a protein also led to distinct specificities, but these were mediated by a spectrum of lower-affinity, ratchet-like docking modes in shallow energy wells, demonstrating a notable plasticity within the same interface. To understand the formation of the new interfaces, we simulated evolutionary trajectories from a machine learning–estimated fitness function that, combined with epistasis analysis, revealed surface residues that act as initiating “seeds” at the earliest stages of an evolutionary path to a high-affinity complex. Our experimental data suggest a hypothesis wherein protein surfaces that have evolved to bind to other proteins have hard-wired biophysical properties conducive for ligand binding, versus other regions of proteins, which have likely been counterselected during evolution to avoid interactions.



Schematic depiction of the initial encounter of a protein-protein complex mediated by seed contacts between phenylalanine residues, which propagates through epistasis to form the complete protein-protein interface.

CONCLUSION: Our protein-protein coevolution platform has shown that there are differences between the surfaces of proteins that have evolved to bind to other proteins versus those of proteins that have not. Indeed, in the system studied, there appears to be a limit on the energetic fitness and affinity of complexes engineered between surfaces of proteins with no evolutionary binding imprint. These data suggest why natural binding sites attract binders: Alternative surfaces have a shallow energy landscape that disfavors tight binding, likely owing to evolutionary counterselection. This may explain, in part, the occurrence of “undruggable” protein surfaces and prompt new biophysically based strategies for overcoming this barrier. □

*Corresponding author. Email: jennl@berkeley.edu (J.L.); kcgarcia@stanford.edu (K.C.G.)
 †These authors contributed equally to this work. Cite this article as A. Yang *et al.*, *Science* 391, eadx6931 (2026). DOI: 10.1126/science.adx6931

PROTEIN INTERACTIONS

Structural ontogeny of protein-protein interactions

Aerin Yang¹†, Hanlun Jiang²†, Kevin M. Jude^{1,3}†, Deniz Akpınaroglu^{4,5}, Stephan Allenspach²†, Alex Jie Li^{4,5}, James Bowden², Carla Patricia Perez⁶, Liu Liu¹, Po-Ssu Huang⁷, Tanja Kortemme^{4,5,8}, Jennifer Listgarten^{2,4,9*}, K. Christopher Garcia^{1,3*}

Understanding how protein binding sites evolve interactions with other proteins could hold clues to targeting “undruggable” surfaces. We used synthetic coevolution to engineer new interactions between naïve surfaces, simulating the de novo formation of protein complexes. We isolated seven distinct structural families of protein Z-domain complexes and found that synthetic complexes explore multiple shallow energy wells through ratchet-like docking modes, whereas complexes formed by natural binding sites converged in a deep energy well with a relatively fixed geometry. Epistasis analysis of a machine learning–estimated fitness landscape revealed “seed” contacts between binding partners that anchored the earliest stages of encounter complex formation. Our results suggest that “silent” surfaces have a shallower energy landscape than natural binding sites, disfavoring tight binding, likely owing to evolutionary counterselection.

How new protein-protein interactions arise, diversify, and evolve distinct specificities remains poorly understood. Yet it is an important problem to better understand, as there appear to be distinctive properties of protein surfaces that have evolved to form protein-protein interfaces versus noninteracting surfaces of proteins (1–3). Natural protein binding sites are usually the most “druggable” sites on a given protein and tend to attract the majority of binders from combinatorial peptide or antibody libraries (4–6). Whether there are biophysical properties that distinguish regions of protein surfaces that have not evolved to bind to a ligand from natural binding sites remains a question. Although it has been proposed that protein binding sites tend to have more exposed hydrophobic amino acids, no definitive rules exist (2, 7, 8). Indeed, protein binding sites are characterized by a vast diversity of chemistries and conformations (9–11). As a result, it remains challenging to identify an unknown binding site solely from inspection of an unbound structure (3).

Protein interfaces often evolve under selective pressures, giving rise to compensatory mutations that confer binding specificity and stability synonymous with particular functions (12–15). Coevolution, characterized by reciprocal changes between interacting partners, reflects structural and functional constraints imposed by evolution and provides insights into the ontogeny of protein-protein interactions (PPIs) (16, 17). Early studies have revealed that residues in close physical proximity at protein interfaces tend to coevolve, exhibiting correlated mutations to maintain structural compatibility and functional integrity (18–20). These coevolutionary signals have been leveraged through computational approaches, such as statistical energy models, to infer residue-residue contacts and identify structural features of

PPIs (16, 21–23). Although such models effectively infer evolutionary constraints from sequence variation, the data on which they are trained may not capture the full range of coevolutionary trajectories in PPI evolution. Consequently, experimental systems can help us to gain mechanistic insights while exploring areas of complex formation not uncovered by nature alone. Such experiments can directly track coevolutionary dynamics and trajectories, including transient intermediates, under controlled selection (13, 14, 24–26).

We previously developed a synthetic coevolution platform that enables bidirectional coevolution of interacting proteins through “library-on-library” selections, providing a synthetic proxy of protein coevolution at a systems level (27). This experiment demonstrated that coevolution can structurally remodel naturally conserved interaction surfaces into a plethora of diverse interface chemistries and fine specificities through amino acid side chain repacking. Notably, however, the remodeled interfaces were constrained into a common docking mode. This observation raised the question of whether these complexes were trapped in a deep energy minimum and thus encumbered with an evolutionary imprint for binding in a stereotyped manner that could not be traversed even with highly diverse libraries. The coevolution platform offers the possibility to address this question by engineering bidirectional interactions between two naïve protein surfaces unencumbered by a natural binding imprint and to study the epistasis and fitness landscapes (28–30) as a proxy for protein-protein evolution.

In this study, we ask whether there are fundamental differences between surfaces of proteins that have evolved to bind to proteins versus surfaces that have not. We have learned that the type of epistasis and structural adaptations that shape specificity and orthogonality differ between natural and nonnatural protein binding sites. Further, in silico epistasis analysis identified initiating “seed” sequences at the earliest stages of an evolutionary path to a high-affinity complex.

Results

Designing de novo interfaces using synthetic coevolution

We carried out a synthetic proxy of protein-protein coevolution by designing interfaces between “silent” surfaces of proteins that have no ligand recognition history. We previously described a synthetic coevolution platform that enables bidirectional sampling of libraries to remodel the existing dimer interfaces between the natural binding surface of a protein Z-domain and its affibody binder (27). We leveraged this coevolution platform to isolate completely synthetic interfaces between protein Z-domain pairs that have no preexisting interaction with one another. Our strategy was to “relocate” the binding interface from the canonical binding region that the Z-domain uses to bind to its natural ligand, the Fe region of immunoglobulin, as well as synthetic proteins, such as affibodies, to a noncanonical site on both protein partners that is not used for ligand binding (Fig. 1A). Protein Z-domains and their engineered binding scaffolds, affibodies, typically interact with target proteins through their H1-H2 helices, which we refer to as the “natural” binding sites (31). However, we simultaneously diversified the noncanonical H2-H3 and H1-H3 regions of two Z-domains to encourage these libraries to converge on one another at sites on the proteins that are not used for binding ligands, which we refer to as “synthetic interfaces.” When accompanied by deep sequencing of the selection rounds, this strategy captures the emergence and coevolutionary trajectories of new protein-protein interactions as they transition from unbiased exploratory states to structurally optimized conformations. This approach does not recapitulate the temporal trajectory

¹Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA. ²Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, USA. ³Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. ⁴The UC Berkeley–UCSF Graduate Program in Bioengineering, University of California, San Francisco, San Francisco, CA, USA. ⁵Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, USA. ⁶Biophysics Program, Stanford University, Stanford, CA, USA. ⁷Department of Bioengineering, Stanford University, Stanford, CA, USA. ⁸Quantitative Biosciences Institute, University of California, San Francisco, San Francisco, CA, USA. ⁹Center for Computational Biology, University of California, Berkeley, Berkeley, CA, USA. *Corresponding author. Email: jenni@berkeley.edu (J.L.); kgarcia@stanford.edu (K.C.G.) †These authors contributed equally to this work

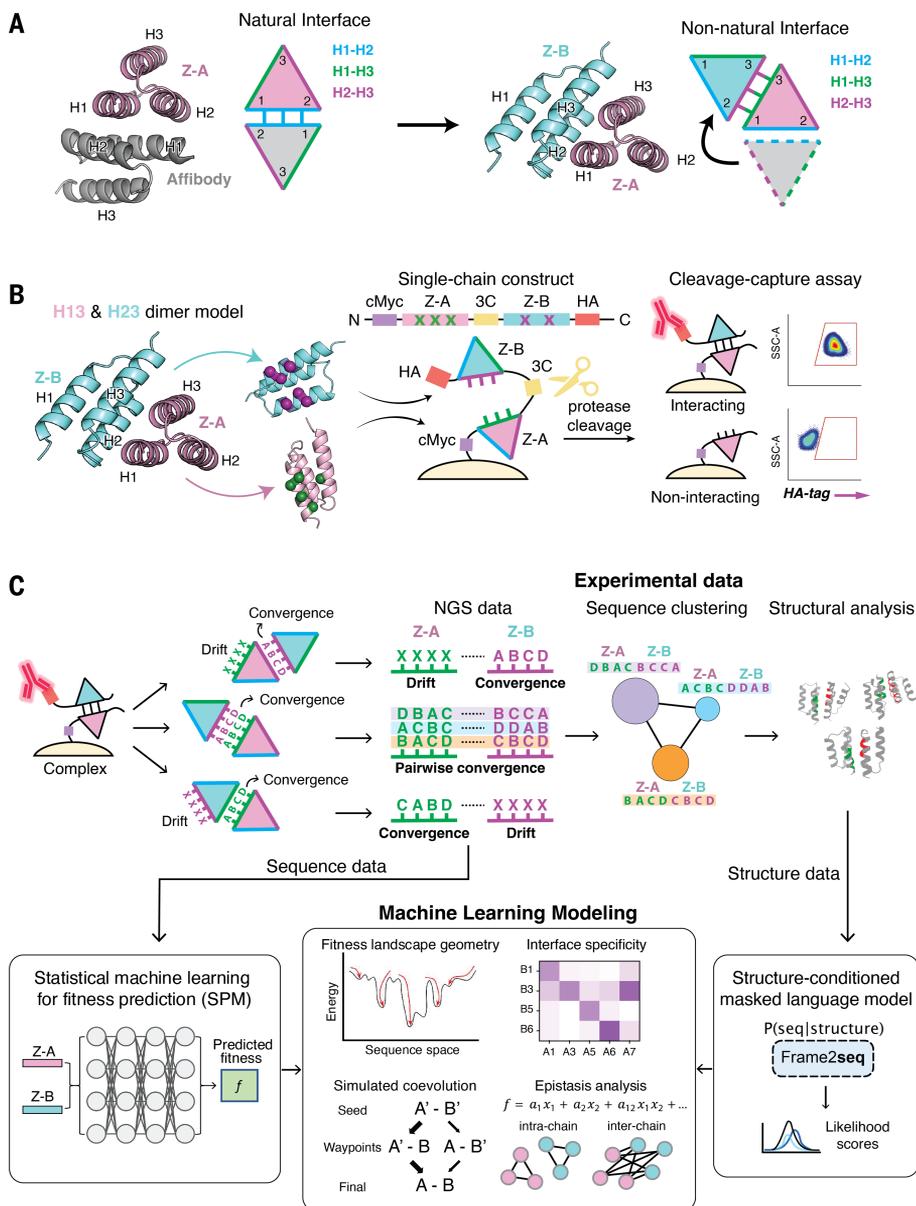


Fig. 1. Synthetic coevolution workflow. (A) Overview of the interface relocation strategy. Canonical binding interfaces (H1-H2) of protein Z-domain and affibody were relocated to noncanonical H2-H3 and H1-H3 surfaces to create new synthetic interfaces. (B) Schematic of the protease-based cleavage-capture assay. Two proteins are displayed on yeast as a single-chain construct connected by a flexible linker containing a 3C protease cleavage site. Upon cleavage, interacting pairs retain fluorescence from the C-terminal HA-tag binding antibody, whereas noninteracting pairs lose the HA signal, enabling differentiation based on binding affinity. Flow cytometry plots were gated using SSC-A (side scatter area, a measure of cell granularity or complexity) and APC-A (HA tag fluorescence readout). (C) The workflow integrates experimental and computational approaches. Yeast display screening followed by NGS is used to track sequence evolution, identify pairwise convergence, and construct SSNs that visualize cluster formation based on sequence relationships. Crystal structures validate these clusters, revealing interface architecture and binding modes. The NGS data are also harnessed by our statistical ML method to model multiround selection data to obtain an estimated fitness landscape, which can be used to uncover landscape geometry through simulating coevolutionary trajectories and to elucidate key protein-protein interactions through epistasis analysis. Frame2seq provides sequence-structure mapping, linking sequence clusters to their structural configurations, using crystal structures as input. Epistasis analysis, derived from these computational models and validated by interface structures, reveals critical residue-residue interactions, providing insights into the evolutionary mechanisms shaping protein-protein interactions.

of natural evolution, as all mutants are contained in the initial library (i.e., at time zero prior to any selection), but stepwise selection with higher stringency can simulate, in some respects, affinity maturation over time.

We used our coevolution platform with a protease-based cleavage-capture assay (Fig. 1B). The assay distinguishes between binding and nonbinding pairs based on fluorescence retention; interacting pairs displayed on yeast retain hemagglutinin (HA) tag fluorescence after 3C protease cleavage, whereas noninteracting pairs lose the signal. The HA tag signal correlates with binding affinities (27), enabling relative comparisons of evolved complexes.

We functionalized the two preselected sites of interaction with diverse libraries. Newly coevolved interfaces were revealed by tracking pairwise next-generation sequencing (NGS) sequence information of interacting pairs, uncovering patterns of convergence indicative of coevolution through epistasis. We considered several possible outcomes (Fig. 1C): (i) Sequence clusters showing linkage between Z-A and Z-B sequences indicate successful library-on-library convergence to form a new bidirectional interface; (ii) sequence convergence in the partner Z-A site paired with sequence drift in the Z-B site indicates that Z-A is adapting to an invariant surface on Z-B and that Z-B sequence drifts because of no selective pressure to adapt to a surface of Z-A, allowing random amino acids; and (iii) drift in the Z-A library paired with convergence in the Z-B library indicates that Z-A has no interactions, whereas Z-B has found an invariant surface on Z-A.

We then leveraged machine learning (ML) and statistical modeling to glean more from our experimental data, helping to strengthen our understanding of sequence-structure relationships and epistatic interactions within them. Specifically, we used our selection probabilistic model (SPM), a statistical ML framework, to extract energy and fitness landscapes from the multiround selection read count data, capturing both round-specific and global fitness landscapes encapsulated in a neural network (materials and methods; fig. S1). Using the global fitness landscape, we simulated *in silico* coevolutionary trajectories, identifying potential evolutionary pathways toward accessible energy minima. In addition, we exploited structural information by using a newly developed inverse folding model, Frame2seq (32), to comprehensively map the sequence-structure compatibility of all 714 sequences in the selected clusters and the six converged structural configurations with solved crystal structures (materials and methods). Lastly, by performing an epistasis analysis of the SPM energy landscape and the Frame2seq sequence-structure compatibility scores, we uncovered key residue interactions that shaped the ontogeny of our protein-protein interactions.

Construction and selection of a coevolutionary library for evolving new protein interfaces

To construct our coevolutionary library, we first screened random Z-domain docking interactions to triage for sterically incompatible surface contours and identified complementary surfaces that might support a new interaction using Rosetta and PatchDock (33). This process involved docking two polyvaline (poly-Val) Z-domains—which show high shape complementarity, with interface areas exceeding 1000 Å²—to identify docking poses free of clashes. Using valines as a generic sidechain representation has been widely used in protein design backbone modeling tasks before amino acid identities are resolved (34, 35). In this case, in docking the two domains, the valines provided a neutral volume in the initial orientation to define the residues in the interface. Selected docked models underwent iterative refinement and energy minimization in Rosetta, yielding a stable, low-energy H2-H3/H1-H3 poly-Val docking model that served as the basis for our library design (Fig. 2A and fig. S2). Based on the docked poly-Val model, we identified key interface residues for randomization, replacing bulky or charged residues to facilitate efficient contacts between proteins at designated library positions (fig. S3). Five positions (8, 11, 14, 15, and 45) on Z-A and six positions (29, 30, 33, 43, 44, and 47) on Z-B were selected for randomization using five hydrophobic amino acids (Met, Phe, Leu, Ile, and Val) to construct the coevolutionary library (Fig. 2A) (36–38). The library was then screened using our coevolution platform, with HA tag fluorescence monitored by flow cytometry to assess enrichment across rounds (Fig. 2B).

Flow cytometry data revealed progressive enrichment of interacting pairs, as indicated by increased HA tag signal retention over successive rounds (Fig. 2C, left). Selection was carried out until the majority of cells showed clear HA tag retention after cleavage, relative to precleavage staining, ensuring that libraries were enriched to comparable functional endpoints. Despite this enrichment, sequence logos indicated substantial diversity within the library, without clear convergence to a single consensus sequence even in the later rounds (Fig. 2C, middle). This suggests that the library preserved multiple viable configurations rather than converging on a single optimal solution. Notably, sequence similarity networks (SSNs) (39) of the top enriched sequences from each round revealed the emergence of distinct clusters in the later rounds, highlighting patterns of convergence at the level of sequence clusters rather than individual residues (Fig. 2C, right). As later analyses indicate, these clusters represent distinct interaction modes evolved at the interface, revealing that the synthetic coevolution platform has captured multiple binding solutions.

Divergent evolution of synthetic interfaces captured through sequence clustering, structures, and fitness landscape

To visualize the relationships among the enriched sequences in the round 5 NGS data, we constructed an SSN (39). Sequence communities within the SSN were merged into single nodes, resulting in a community map comprising seven distinct clusters (Fig. 3A). Although the overall amino acid composition across all 11 library positions remained diverse, sequence logos for each cluster showed patterns of convergence, suggesting distinct coevolutionary pathways to the different clusters. Notably, we only saw clusters with convergence of both interfaces, suggestive that the partners had “homed” in on one another through their library patches, albeit with imperfect structural alignments of the libraries (discussed in “Structural insights into specificity and cross-reactivity between clusters”) (Fig. 1C). To further demonstrate the specificity and orthogonality of interactions, we used a Circos plot (40) to represent pairwise relationships between Z-A and Z-B proteins within the round 5 NGS data (Fig. 3B). The plot revealed minimal cross-reactivity (orthogonality) between clusters, with most sequences paired within their respective clusters.

To explore the structural characteristics of coevolved complexes, we selected a representative of the most enriched sequences from each cluster and validated their binding using the cleavage-capture assay, confirming strong binding (Fig. 3C). For six of the seven clusters, we solved crystal structures of one coevolved complex, with resolution ranging from 1.43 to 2.85 Å (table S1). Structural analysis revealed considerable diversity in the docking orientations across the complexes while maintaining unidirectional polarity (i.e., we did not see reverse docking orientations) analogous to the hands of a clock. Docking angles varied by up to 91.1° between A5B5 and A7B7 complexes, and none of the coevolved docking poses resembled the initial docked poly-Val model (Fig. 3D and fig. S4). This wide diversity in ratchet-like docking orientations contrasts with our previous study (27), where highly constrained docking modes were observed in coevolved complexes involving the natural interface of the Z-domain and its affibody binder (Fig. 3E). Interresidue distances across the natural interface complexes were highly conserved, whereas those of the synthetic interface complexes varied substantially (fig. S5). The buried surface areas (BSAs) across the coevolved interfaces ranged from 1080 to 1475 Å², which are on the smaller side for PPIs but remain consistent with those of previous structural studies (11, 41). To assess interface packing, we calculated shape complementarity (sc) (42) packstat scores (43) and analyzed cold spots (44–46) across all solved natural and synthetic interface complexes (fig. S6). We identified a single cold spot (44) in the parent structure of the natural interface (fig. S6A), whereas no cold spots were detected in any of the coevolved natural or synthetic interfaces, indicating that the coevolved interfaces are well packed.

Despite the library composition being limited to generally hydrophobic residues, the diversity in docking orientations of the synthetic interface clusters, spanning a range of approximately 90° in 15° increments, suggests that distinct fitness landscapes were sampled during coevolution. To elucidate characteristics of the overall fitness landscape, we used a simple *in silico* coevolution simulation (one mutation at a time and only improving the fitness), generating 10 million trajectories for each interface, using our learned fitness function from SPM (materials and methods). For each sequence, we computed its accessibility, which is defined as the fraction of all trajectories that landed at that sequence. Highly accessible sequences represent energy wells in the energy landscape where coevolutionary pathways converged [energy is the negative fitness (47)]. We noticed a marked difference in the number and depth of energy wells between the natural and synthetic interfaces (Fig. 3F). For the natural interface, the energy landscape was dominated by a single highly accessible well containing all strongly binding sequences. By contrast, the energy landscape of the synthetic interface contained a multiplicity of comparably accessible wells that each contained strongly binding sequences (Fig. 3F and fig. S7). These differences, namely, those of well depths and accessibilities between natural and synthetic, were substantiated as statistically significant through a bootstrap-based statistical test (fig. S8; materials and methods), with *P* values of 4.7×10^{-310} and 2.9×10^{-147} , respectively. We additionally experimentally characterized the geometries of the energy landscapes as follows. We measured dissociation constants (K_d) by surface plasmon resonance (SPR) for a subset of the variants in the energy landscape for each of synthetic and natural interfaces. Specifically, we leveraged K_d measurements for six variants in the synthetic interface experiments appearing in the specificity matrix (fig. S9), which happen to be among the 20 most accessible appearing in our energy landscape plot. Their measured K_d values ranged from 0.4 to 8.8 μM, which corresponds to an energy gap of 1.8 kcal/mol. For the natural interface, we obtained SPR measurements for the variant corresponding to the deepest energy well and four variants that lie at the bottom of four other wells in the natural landscape. The K_d value of the variant corresponding to the deepest energy well was 2.4 nM (27), and for the

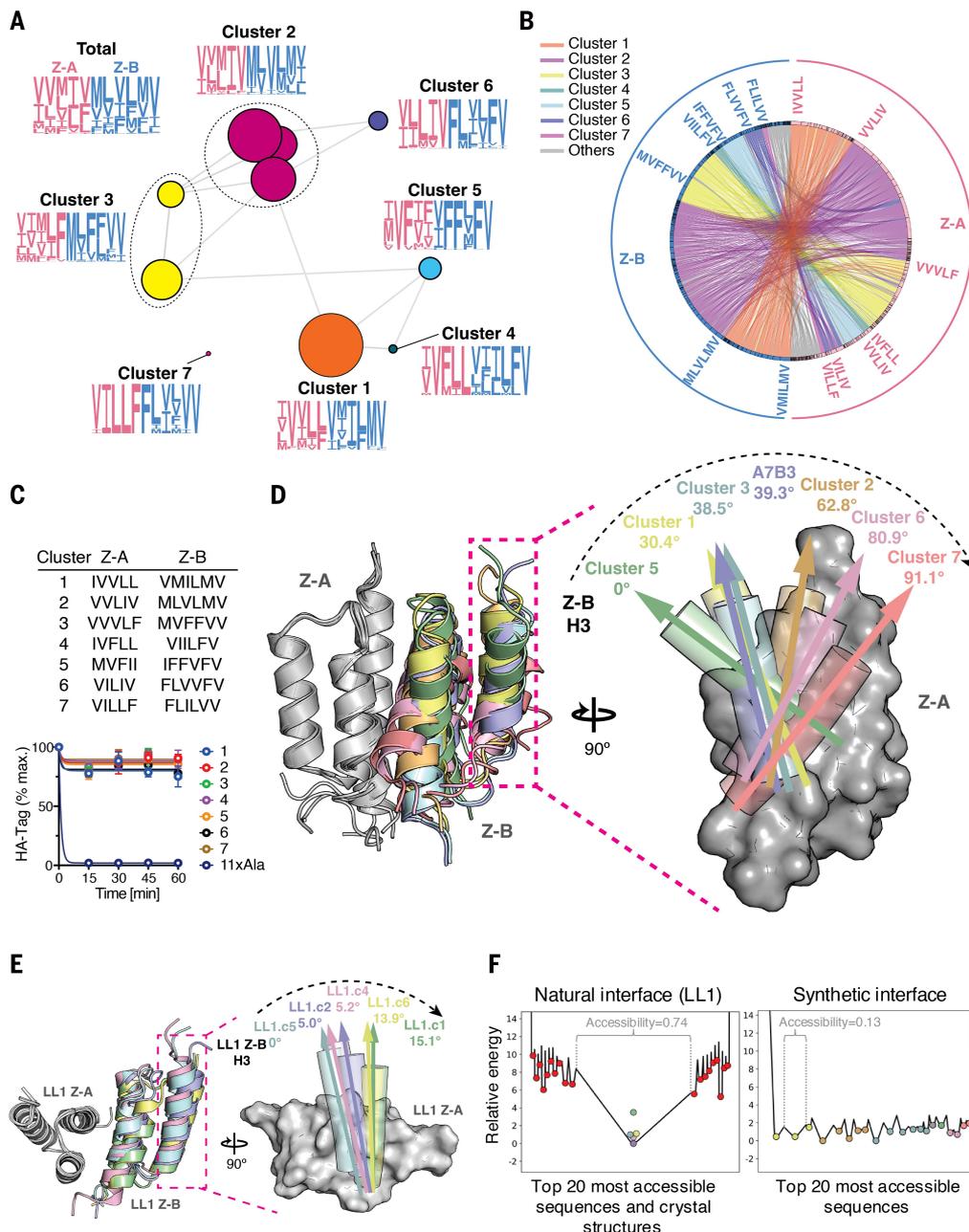


Fig. 3. Sequence clustering and fitness landscape of synthetic interfaces. (A) A community map of SSN of round 5 sequences. Clustered communities are merged into single nodes to form a community map. Seven distinct clusters were identified, each exhibiting distinct patterns of convergence, as illustrated in the sequence logos. (B) A Circos plot (40) representing pairwise relationships between Z-A and Z-B proteins in round 5 NGS data. Each pair is normalized to have equal area, providing a visual representation of the approximate specificity or cross-reactivity of each sequence. The plot highlights high orthogonality, with sequences primarily pairing within their respective clusters and minimal cross-cluster interactions. (C) Validation of representative interactions. (Top) A table of representative pairs from each cluster. (Bottom) On-yeast cleavage-capture assay confirms strong binding affinities for all tested pairs, validating functional interactions. Data are mean \pm SD; $n = 3$ independent replicates. (D) Structural diversity in synthetic interface docking orientations. (Left) Superimposed crystal structures of synthetic interface complexes. (Right) Docking angle variations of Z-B helix 3 relative to the cluster 5 structure, highlighting the diversity in docking geometries among clusters. (E) Conserved docking orientations in natural interface complexes. Structures from a previous study (27) reveal that natural interface complexes of the Z-domain and its affibody binder exhibit relatively consistent docking angles, in contrast to the synthetic interface complexes in (D). (F) The relative energy landscapes of the 20 most accessible wells for the natural interface (LL1 library) and the synthetic interface. Energy is defined as the negative fitness (47), and, for ease of comparison of the interfaces, for each interface, we computed a relative energy for each sequence by subtracting the energy of the strongest-binding sequence. We did so because relative energies are comparable between the two plots (see supplementary text section C for details), whereas energies are not. The width of each energy well corresponds to the accessibility (see main text and materials and methods) of that well. In each subplot, the most accessible energy well is marked by the curly bracket labeled with its accessibility. For the natural interface, red dots denote the 20 most accessible sequences that are not crystal structures, whereas colored dots correspond to the crystal structures in (E), all falling into the same well, consistent with their conserved docking modes. For the synthetic interface, dots are colored according to the colors of the structures in (D) based on their corresponding clusters. To visualize the geometry of each landscape, we chose a random ordering of the 20 most accessible wells, with the exception of the natural interface, which always centers the deepest energy well, using a random ordering for the rest.

other four sequences, it ranged from 6.7 to >100 μM (i.e., beyond the SPR-measurable limit of 100 μM) (fig. S9D), for which the smallest energy gap between the deepest well variant and the other variants was 4.7 kcal/mol (from the 6.7 μM variant). Notably, this natural interface landscape energy gap (conservatively estimated because it is the smallest one) is 2.6 times larger than the measured energy gaps in the synthetic interface.

Structural insights into specificity and cross-reactivity between clusters

To further elucidate the specificity and cross-reactivity of coevolved complexes, we analyzed the structural and sequence features of representative pairs from each cluster. Specificity between the seven representative pairs was validated by measuring relative binding affinities of all possible combinations of Z-A and Z-B proteins (Fig. 4A). The resulting specificity matrix revealed strong intracuster specificity, with cognate pairs occupying the diagonal positions. Some instances of cross-reactivity were observed, including A4B1 and A7B3 mixed complexes, which formed between noncognate pairs. That clusters can cross-react, despite evolving distinct sequence and structural solutions, suggests that the different docking modes are separated by shallow energetic barriers, consistent with the fitness analysis (48). To complement these analyses, we measured solution affinities (K_d) by SPR for all combinations in the specificity matrix. Cognate pairs as well as the noted cross-reactive pairs (A4B1 and A7B3) bound in the sub- to low-micromolar range (0.3 to 1.7 μM). Notably, the solution K_d values correlated strongly with the on-yeast cleavage-capture HA tag readout [Spearman rank correlation coefficient (r) = 0.700, Pearson correlation coefficient (r) = 0.868, coefficient of determination (R^2) = 0.754; fig. S9], confirming that the specificity matrix accurately reflects binding in solution. We also examined whether biophysical parameters aligned with these affinities (11); however, common metrics, such as BSA, sc , and the hydrophobic fraction of buried surface area (% hydrophobic ΔSASA), showed no correlation with K_d , underscoring that structural features alone are insufficient predictors of affinity in our synthetic interfaces (fig. S10).

The experimental cross-reactivity matrix for the seven representative sequence pairs implies that an unexpectedly high degree of orthogonality arises from the relatively limited choice of hydrophobic residues at each library position. To probe the generality of this observation beyond individual sequences, we sought to comprehensively map all sequences in the clusters to the six crystal structures. To do so, we used Frame2seq (32) to quantify sequence-structure compatibility in silico (Fig. 4B; materials and methods). We provided the experimentally solved crystal structure backbones and sequences outside the library positions as inputs to the model and computed Frame2seq model scores to estimate sequence-structure compatibility for the library residues at the interface (eq. S1; materials and methods;). We analyzed (i) chain pairing specificity (Fig. 4C) and (ii) the compatibility of all selected sequences in each cluster with each of the six crystal structures (Fig. 4D). To assess chain pairing specificity, we scored all possible Z-A and Z-B combinations with cognate (diagonal) and noncognate (off-diagonal) chain pairings (Fig. 4C). The resulting chain pairing confusion matrix showed a strong correlation to the experimentally validated specificity matrix (Spearman r = -0.746). This level of predictive accuracy of Frame2seq for specificity prediction was notable given the relatively shallow fitness landscapes between clusters and is encouraging for using such models for mapping sequence-structure space. To further assess the generalizability of Frame2seq, we benchmarked it against a dataset of experimentally measured interface mutations used in (49), where the model successfully distinguished destabilizing from stabilizing mutations (fig. S11), consistent with its predictive accuracy on our synthetic complexes. To assess sequence-structure mapping, with each of the six experimentally solved structures representing

these clusters (Fig. 4D). Cognate sequence-structure pairings were consistently scored by Frame2seq to form more favorable (lower scoring) interfaces. These results suggest that the sequences in each of the identified clusters indeed map well to the six representative crystal structures.

We examined the high-resolution crystal structures of coevolved complexes from each cluster to understand the basis of both specificity and cross-reactivity. The structures revealed a wide range of interface packing solutions, with converged residue interactions driving specificity within clusters (Fig. 4E). Converged residues, marked with asterisks on the sequence logos, were highlighted in stick representation on the structures. These residues cluster closely within the interfaces, emphasizing that pairwise convergence observed in NGS data correlates strongly with cluster-specific interactions and distinct structural determinants that define docking orientations.

Notably, beyond intracuster specificity, we also observed intercluster cross-reactivity that can result in dual or multispecificity by a single binding partner. To assess how common these cross-reactivities are, we quantified intercluster cross-reactivity among the sequences from the SSN clusters (Fig. 3A). We found that 19.9% of Z-A sequences exhibited at least one intercluster cross-reactivity, whereas only 3.4% of Z-B sequences did so, indicating that cross-reactivity occurs primarily on the A chain (fig. S12). This is exemplified by the A7 protein, which can bind strongly with both B3 and B7 with comparable affinities (K_d = 487 nM for A7B7, 551 nM for A7B3) (Fig. 4F and fig. S9). Structural analysis revealed that A7 adopts different docking modes depending on its partner. In the A7B3 complex, the Z-B subunit is rotated by 53.9° relative to the A7B7 complex, and the H1 helix of A7 shifts by 4.1 Å (Fig. 4F and fig. S13). This demonstrates how sequence convergence within clusters supports intracuster specificity, whereas intercluster cross-reactivity emerges from docking plasticity, allowing a shared partner to navigate distinct energy wells through structural adaptation. Notably, cross-reactivity was also directional, with A7B3 forming but another possible cross-reactive pair A3B7 failing to interact owing to a steric clash caused by a M29^BF substitution in Z-B (fig. S13). Notably, position 29^B strongly converges to Met in cluster 3 and Phe in cluster 7 (Fig. 4E), suggesting that these residues contribute to cluster-specific docking preferences and impose structural constraints on intercluster cross-reactivity.

Structural and epistatic signatures underlying synthetic and natural protein-protein interfaces

Our hypothesis in the design of the coevolution experiment was that the library patches would converge on one another through interchain epistasis. To ask how structural convergence occurred, we first curated all atomic contacts <4 Å between subunits in all of our crystal structures (Fig. 5A and table S2). We then classified randomized positions as “library” and constant positions as “framework,” denoting contacts as library-on-library “LL” ($A_{\text{lib}}-B_{\text{lib}}$), library-on-framework “LF” ($A_{\text{lib}}-B_{\text{frame}}$, $A_{\text{frame}}-B_{\text{lib}}$), or framework-on-framework “FF” ($A_{\text{frame}}-B_{\text{frame}}$). The cores of all interfaces are dominated by LL contacts, which presumably act as “hotspots” to initiate complex formation at the early stages of selection. The centrally located LL patches are not perfectly structurally coincident, thus LF contacts occur at the peripheries of the misaligned LL regions, and FF interfaces occur adventitiously at the outermost edges of the interfaces.

We sought to quantify the distinct patterns of structural interactions in the natural versus synthetic interfaces by comparing their contact ratios, defined as the number of LL contacts ($A_{\text{lib}}-B_{\text{lib}}$) divided by the number of LF contacts ($A_{\text{lib}}-B_{\text{frame}}$ and $A_{\text{frame}}-B_{\text{lib}}$) in the crystal structures. Synthetic interfaces spanned a wide range of contact ratios, with an average of 0.61, whereas the natural interfaces had a narrower range of contact ratios and a lower average: 0.22 for the LL1 library and 0.33 for the LL2 library (Fig. 5B, left). We hypothesized that the LL interactions might be related to interchain epistasis, whereas LF

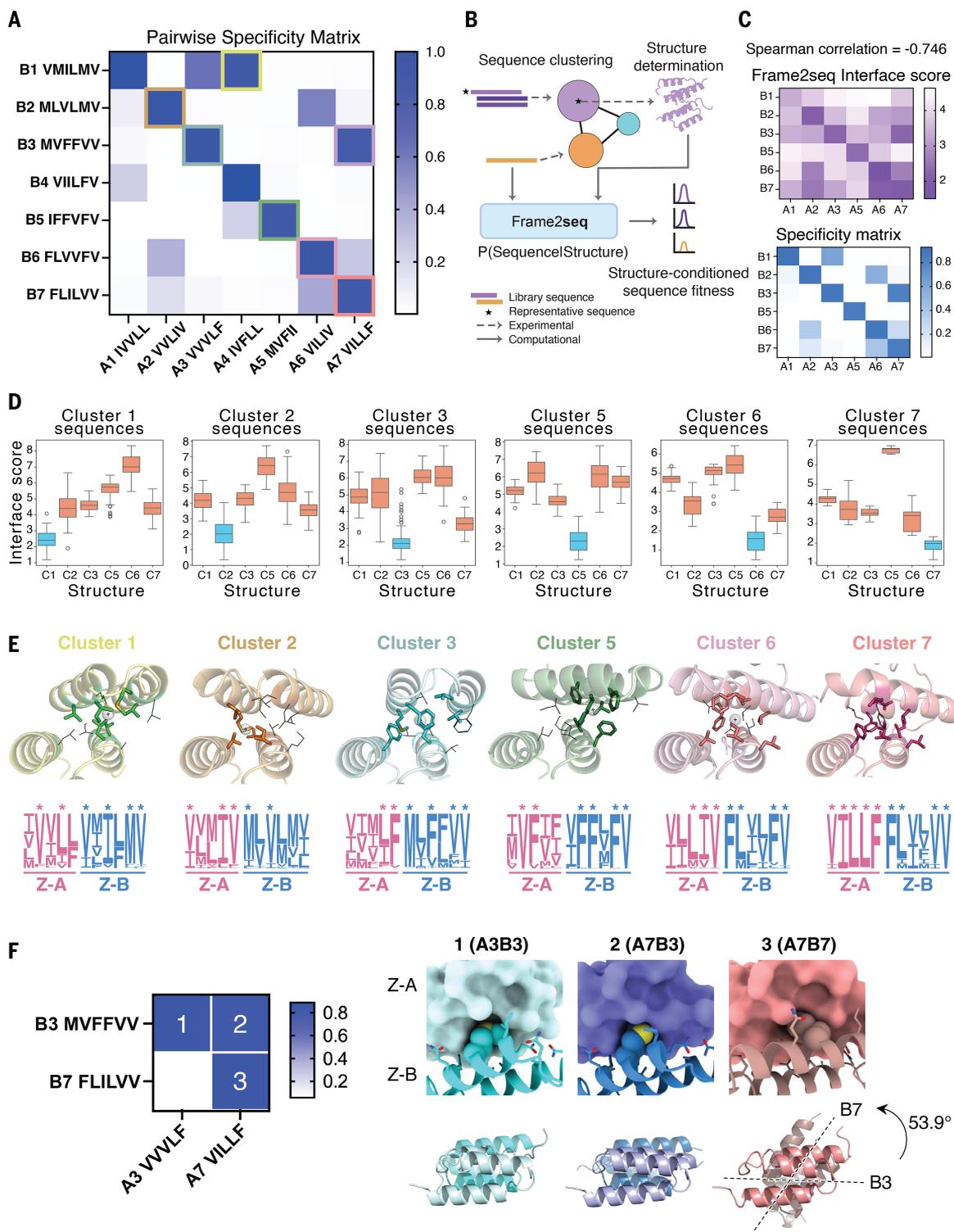


Fig. 4. Structural insights into specificity and orthogonality between clusters. (A) Specificity matrix of all possible Z-A and Z-B combinations of representative pairs in Fig. 3C. Binding affinities measured by on-yeast cleavage-capture assay were normalized to the highest affinity (A4B4 complex) and noninteracting control (11xAla mutant). The matrix shows high intracluster specificity and some intercluster cross-reactivity (e.g., A4B1, A7B3). Data are mean of $n = 3$ independent replicates. (B) Workflow for scoring sequence-structure compatibility using Frame2seq. All library sequences were scored for compatibility with their representative structure. (C) Predicted Frame2seq interface score matrix (top) and experimentally determined specificity matrix (bottom) of representative sequences (Spearman $r = -0.746$). Frame2seq scores are negative log likelihoods (lower scores indicate higher fitness). (D) Frame2seq interface scores of round 5 SSN cluster sequences on all representative crystal structures (see materials and methods for details on representative structure selection). The correct (cognate) sequence-structure pairings are in blue, and the incorrect (nongonate) pairings are in orange. (E) Interface structures reveal converged residue interactions. (Top) Interface structures of coevolved complexes highlight converged residues in stick representation, with other library residues shown as lines. The center of mass (COM) of the converged residues is depicted as a gray sphere at the interface, illustrating their clustering in proximity. Each structure exhibits distinct COM positions, reflecting different interaction and docking geometries across clusters. (Bottom) Sequence logos of each cluster illustrate residue

convergence, with converged residues marked by asterisks. These data demonstrate how pairwise residue convergence underpins cluster-specific interactions and correlates with distinct docking orientations, highlighting the structural determinants of specificity within coevolved interfaces. **(F)** (Left) Pairwise specificity matrix highlighting biased cross-reactivity between cluster 3 and 7 pairs. (Right) A7 exhibits dual specificity by binding both B3 and B7, whereas A3 binds only B3. The interaction intensity in the left matrix is derived from the specificity matrix in (A). Interfaces and top views of the A3B3 (pale cyan and cyan), A7B3 (slate and sky blue), and A7B7 (salmon and dark salmon) crystal structures. Z-A subunits are shown as surface representations and Z-B subunits as cartoons, with residue 29^B side chain atoms shown as spheres. Top views reveal distinct docking geometries: In the A7B7 complex, the Z-B subunit is rotated by 53.9° relative to the A7B3 complex. Models of putative A3B7, which fails to form a stable complex, are presented in fig. S13.

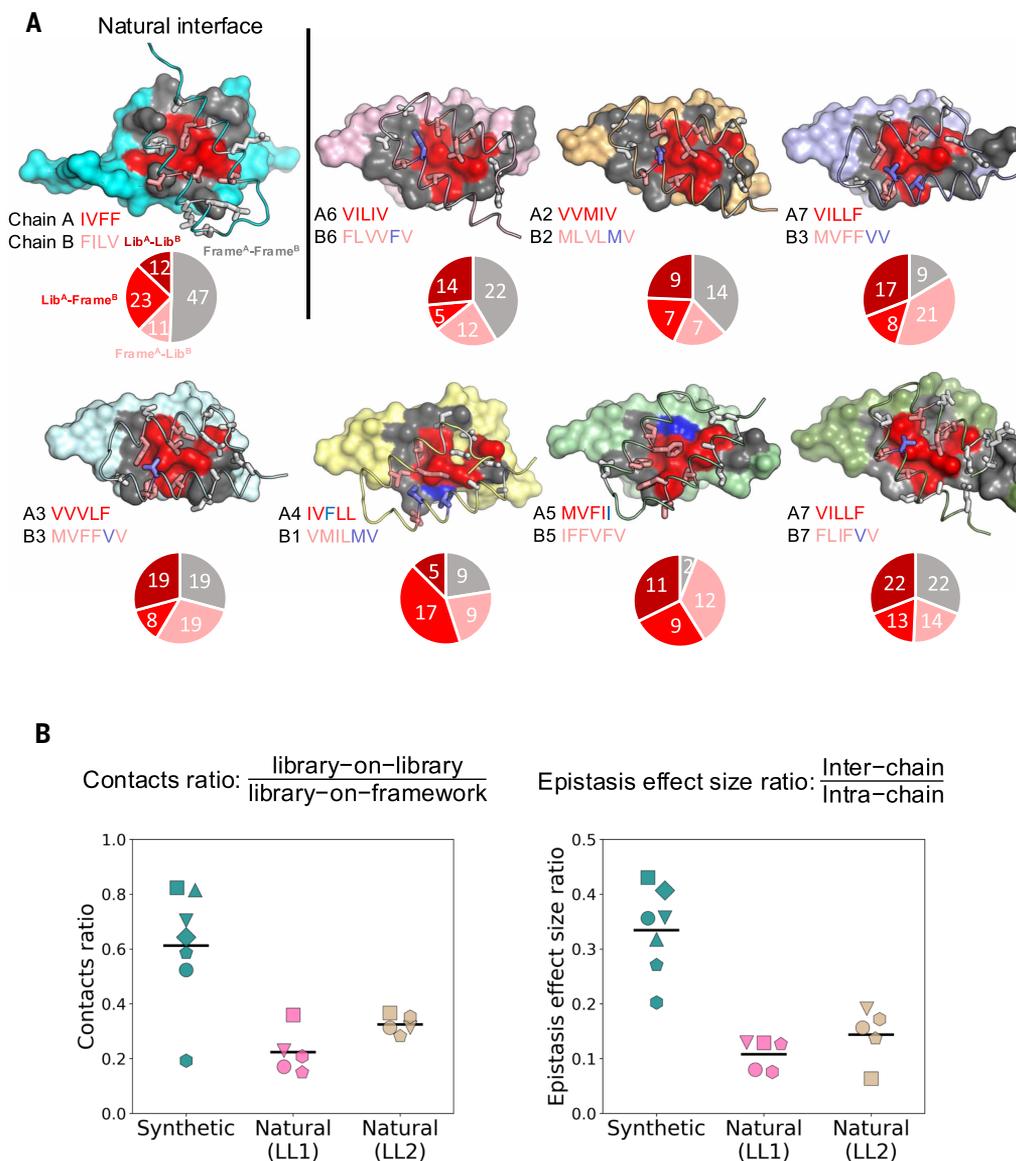


Fig. 5. Structural parsing of interface composition and analysis of epistasis. **(A)** Contacts between Z-domain chains. For each structure, chain A is shown as a surface representation, and chain B is shown as a cartoon with contacting residue side chains shown as sticks. Noncontacting library residues are blue or slate, contacting framework residues are gray or white, chain A contacting library residues are red, and chain B contacting library residues are salmon. The noncontacting helix of chain B is omitted for clarity. Below each structure is a pie chart of atomic contacts between helical residues: A framework–B framework (gray), A framework–B library (salmon), A library–B framework (red), A library–B library (dark red). A representative natural interface Z-domain–affibody pair (LL2.c22) is shown in cyan. **(B)** Dot plots of contact ratios (library-on-library contacts over library-on-framework contacts) and SPM epistasis effect size ratios (interchain epistasis effect sizes over intrachain epistasis effect sizes) for the crystal structures. For each type of interface, the black horizontal bar marks the average contacts ratio or epistasis effect size ratio of the corresponding crystal structures. Each crystal structure uses the same markers in the two plots.

interactions might be related to intrachain epistasis. Indeed, we found that a higher ratio of interchain over intrachain epistasis effect sizes tended to correspond to a higher ratio of LL to LF contacts (Fig. 5B, right, and fig. S14). Notably, the de novo complex formation has a higher proportion of LL contacts compared with that of the natural complex,

suggesting that LL contacts may be primary drivers of de novo complex formation. Presumably, this difference arises from the natural protein partners being trapped in the deep energy well of the natural docking mode; within that well, the primary role of epistasis is to optimize contacts in the existing interface rather than to explore new interfaces.

Epistatic dissection of coevolution in synthetic and natural interfaces

We next sought a more detailed understanding of how epistasis and structure play out during coevolution. Physical interactions could form in part through intermolecular (ZA-ZB) epistasis of library amino acids on their respective surfaces and also through intramolecular epistasis within each binding partner (ZA or ZB). Intramolecular epistasis could preorganize a binding surface independently of a partner to make it more likely to bind to any partner. To determine which mechanisms may drive different stages of coevolution, we leveraged our epistasis analysis to reveal critical interactions in the overall SPM-predicted fitness landscape as well as those in the individual SPM-predicted fitness landscapes associated with each selection round (see materials and methods, “Extraction of epistasis importance from SPM-predicted fitness landscape”). We also applied a similar analysis to the Frame2Seq sequence-structure compatibility scores.

First, leveraging the SPM fitness, we found a strong correlation between pairwise epistasis importance and physical proximity (Spearman $r = -0.550$; fig. S15). Epistasis “importance” refers to the maximum amount that a given interaction can change the fitness. The weaker correlation for interchain pairs likely arises from diverse docking modes in the crystal structures, as some pairwise epistasis may be important for a particular docking mode. Residues involved in the strongest pairwise intrachain epistatic interactions tended to lie on the same helix at the binding interface (Fig. 6A).

Next, we inspected how contributions of intra- and interchain interactions likely changed over selection rounds by leveraging our round-specific epistasis analysis as a proxy. As expected, we found that both intra- and interchain epistasis increased with increasing selection strength (i.e., increasing round number) (fig. S16). The relative contribution of interchain over intrachain epistasis importance decreased with increasing selection strength (Fig. 6B), suggesting that after establishing binding contact in the early rounds, each chain may then individually focus more on fine-tuning of its own energetics. By contrast, when performing the same analysis on the original Z-domain-affibody interface, the relative contributions did not systematically decrease (Fig. 6B), presumably because binding contact did not need to be established in the first place.

We additionally performed epistasis analysis on the sequence-structure compatibility scores derived from Frame2seq using each of the six crystal structures (materials and methods, “Extraction of structure-conditioned epistasis importance using Frame2seq”). We found that many of the inferred epistatic interactions were conserved (involving equivalent residues) across all six structures, despite the different docking modes (Fig. 6C). Furthermore, several of the identified interactions were also seen in the SPM global fitness epistasis analysis (fig. S17). These results suggest that a set of key interactions are important in the formation of the interface and that the different docking modes “pivot” around these core interactions.

Seed sequences in the coevolutionary paths

We wondered whether there were primordial “seed” contacts that could serve as the earliest sentinels of protein binding emerging from the formerly silent surfaces in the selections. Moreover, such key interactions could be preconditions that favor an evolutionary path toward a specific binding conformation. To attempt to extract the identity of seed sequences, we analyzed our simulated coevolutionary trajectories and curated potential seed sequences as those that are weak binders (with SPM-predicted fitness between 18.35 and 22.07) with relatively high “exclusivity” and “contribution” toward their most likely energy well (the well that their trajectories fall into most frequently; we considered only wells that correspond to sequences in Fig. 4A whose binding affinity had been experimentally measured with normalized HA tag binding affinity >0.8). Exclusivity of a sequence for a well refers to the fraction of trajectories from that sequence that end in that well; contribution of a sequence to a well is

defined as the fraction of all trajectories ending in that well that are accounted for by that sequence (materials and methods).

This curation yielded plausible seed sequences for sequences A7B3 and A1B1. From these seed sequences, we selected one sequence with high exclusivity and relatively high contribution for each of A7B3 and A1B1 for further analysis (fig. S18). We found only one energetically favorable coevolutionary path to A7B3 from its selected seed (Fig. 7A and fig. S19). Our epistasis analysis identified F45^A-F33^B as the highest-ranking interchain pairwise epistasis effect size along this path, suggesting its pivotal role in progression from weak to strong binding partners. This interaction is physically located in the core of the protein interface (Fig. 7B). In contrast to the case of A7B3, for A1B1, we found multiple viable coevolutionary paths from the selected seed (Fig. 7C and fig. S18B). For all the sequences along these paths, the interaction V14^A-V47^B ranked as the highest or second highest in interchain pairwise epistasis effect size, suggesting its pivotal role in binding evolution; residues in this interaction pair are in close physical contact with each other (Fig. 7D). Altogether, these results suggest that key interactions not only help to usher the evolutionary path toward a specific binding conformation from a seed but are also maintained along the path. Moreover, distinct key interactions may drive the differences between docking modes.

Discussion

We have used a coevolution platform to try to understand the differences between surfaces of proteins that have evolved to bind to other proteins versus those that have not. Our strategy was to simulate the ontogeny of new protein-protein interactions in vitro between “silent” surfaces of proteins with no natural history of protein-protein interactions and then deconstruct this process. By integrating experimental coevolution and statistical ML, this platform has revealed how structural and epistatic principles, including “seed” interactions, shape binding specificity, and adaptability. Synthetic complexes adopted a range of docking modes, distinguished by subtly different interfaces, by sampling shallow, sawtooth energetic landscapes. This was not seen during coevolution of a complex mediated by natural Z-domain binding surfaces, which appeared to reside within a deep energy well (27).

Our results offer insight into why surfaces of proteins that have not evolved to engage ligands are notoriously difficult to drug (50). Natural protein binding sites have been evolutionarily refined over millennia, appearing to possess structural and energetic imprints that make them conducive for binding and as drug targets (6, 51). It is well established that combinatorial libraries undergoing unbiased selections, such as peptide and antibody phage libraries, frequently converge on the natural binding sites of proteins, which serve as the most “druggable” sites for both small molecules and protein therapeutics (6). By contrast, regions of protein surfaces that do not engage ligands were likely evolutionarily “counterselected” to avoid spurious interactions with nonspecific proteins, and this likely contributes to lack of “druggability.” Targeting nonbinding regions of proteins may be limited by a relative lack of evolutionary fitness that results in a dearth of structural features conducive to binding (1, 2, 6). Consequently, non-binding surfaces of proteins may be constrained to shallow energetic landscapes that have evolved to avoid nonspecific interactions with irrelevant molecules. Our synthetic coevolution experiment overcame this limitation through molecular diversity, but even so, the resulting complexes exhibited a sawtooth energetic landscape that explored parallel evolutionary pathways to diverse structural solutions. This plasticity may have been further facilitated by the hydrophobic nature of our designed library, which promotes chemically compatible packing across multiple configurations, as shown by the lack of cold spots in the synthetic interfaces despite the shallow energy landscapes (38, 44). Underscoring this shallow energetic landscape, we observed that docking modes can interconvert in a ratchet-like fashion through

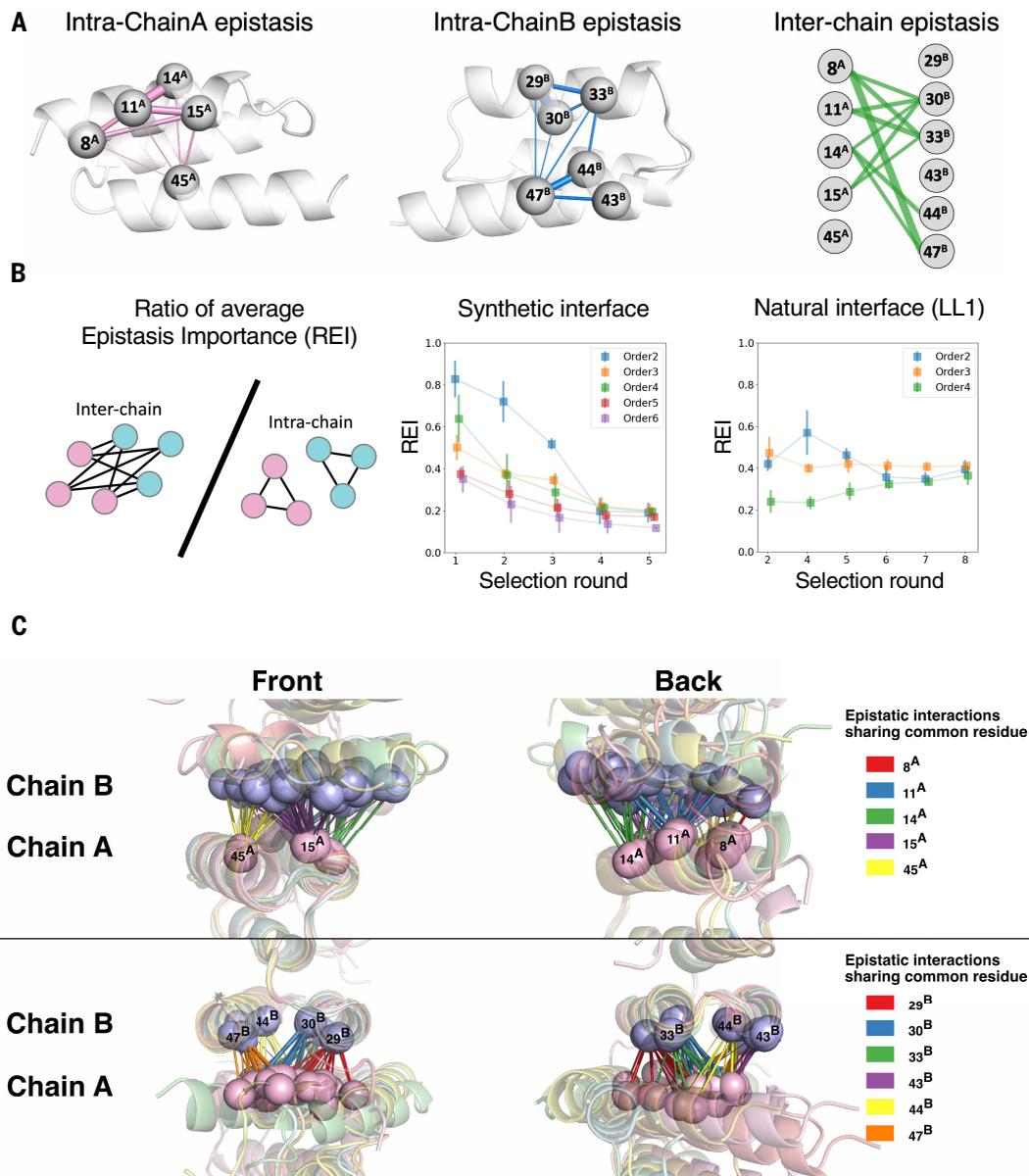


Fig. 6. Epistatic contributions between coevolved pairs. (A) (Left and center) Visualization of intrachain pairwise SPM epistasis overlaid onto the crystal structure of A5B5. Widths of the bonds connecting the library residues are proportional to the strength of the corresponding epistasis importance. (Right) Top 10 interchain pairwise SPM epistasis importance. The widths of the lines are proportional to the corresponding epistasis importance. (B) (Left) Illustration of the ratio of average epistasis importance (REI). It was computed as the ratio of the average interchain SPM epistasis importance per term over the average intrachain SPM epistasis importance per term for each order of interaction. (Middle and right) Comparison of REI for each round of selection experiments of synthetic and natural interfaces. (C) Pairwise epistasis terms per cluster structure, overlaid and aligned on Z-A (above) and Z-B (below) when viewed from the front (left) and back (right) of the complex. Spheres represent library residues, and bars represent epistatic interactions, where thickness represents epistasis strength as computed by Frame2seq, and color represents the residue shared among all such interactions (which is also depicted on the legend in the right).

only a few amino acid mutations, demonstrating a remarkable potential for plasticity in molecular recognition.

A limitation of our coevolution experiment is that we used a biased library restricted to hydrophobic amino acids. We designed our library to be within the experimental limits of diversity of yeast display, which is about 10^9 . We chose 11 sites to cover a surface area that led to binding interfaces on the smaller side for PPI (~ 1080 to 1475 \AA^2) but still common (11, 41) and allowed complete sampling of the theoretical diversity of the library (3.6×10^8). We also wanted the library to be similar in composition to our previous library that we used for remodeling the protein Z-affibody interface (27) so we could make comparisons to the

synthetic interfaces in the current manuscript. Indeed, several of our results are consistent with those of previous studies on evolution of PPI based on analysis of all 20 natural amino acids. For example, we found that dual-binding states [i.e., promiscuous intermediates (13)] frequently occurred along shortest viable mutational paths between coevolved crystal backbones (as per our SPM-estimated fitness); this result suggests that the biased libraries are capable of mimicking characteristics of natural repertoires at a systems level. We also did not see correlations between measured biophysical parameters (BSA, sc, and % hydrophobic Δ SASA) and $pK_a[-\log(K_d)]$, as measured by SPR for our synthetic interfaces, as previously shown in (11). Consistent with

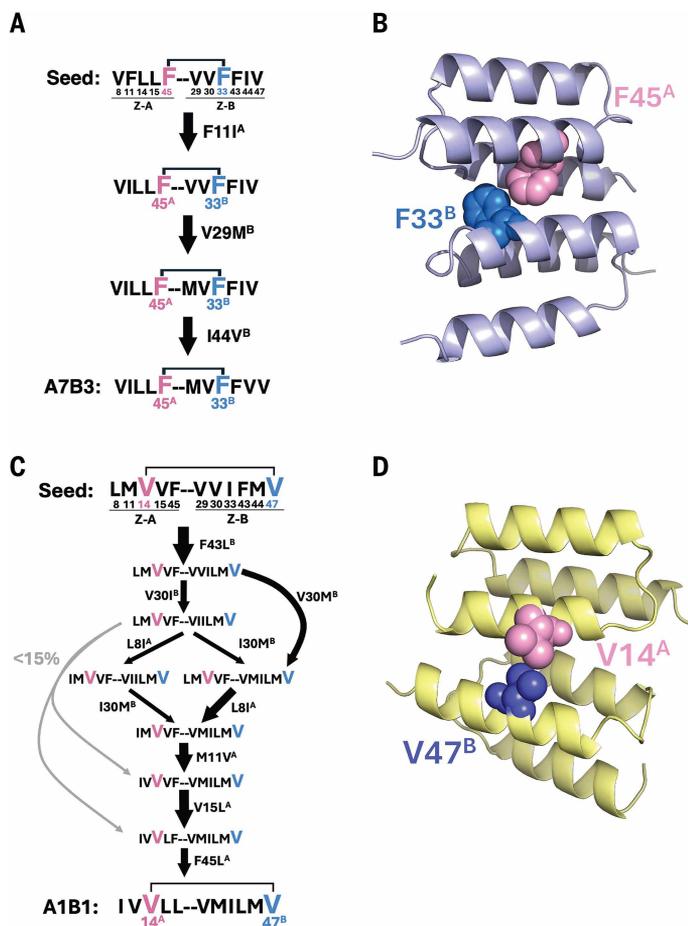


Fig. 7. Seed sequences initiate coevolutionary trajectories. (A) Illustration of the only energetically favorable coevolutionary path from the identified seed (VFLLFVVFIV) to A7B3. The conservation of the top-ranking interchain epistasis between F45^A-F33^B is highlighted. (B) Visualization of the conserved top-ranking interchain epistasis F45^A-F33^B on the crystal structure. (C) Illustration of energetically favorable coevolutionary paths from the identified seed (LMMVFVIFMV) to A1B1. The widths of the arrows are proportional to the number of simulated trajectories visiting the corresponding paths. Gray arrows represent minor paths that were collectively visited by less than 15% of all the simulated coevolutionary trajectories. Conserved top-ranking pairwise interchain epistasis V14^A-V47^B is highlighted in each sequence in the paths. (D) Conservation of top-ranking interchain epistasis visualized on the structural model of A1B1: V14^A-V47^B is the highest- or the second highest-ranking interchain epistasis effect size for the seed, all the waypoints, and A1B1. The structural model of A1B1 was built by making the F14^AV point mutation of the crystal structure A4B1.

well-packed natural PPI using all 20 amino acids, we found that the synthetic interfaces were well packed, lacking cold spots (44, 45), showing that even our limited library is capable of evolving natural-like interfaces.

Concordant with previous studies showing the importance of epistasis in natural protein evolution (15, 46, 52), we also found that epistasis plays a critical role in the evolution of both synthetic and natural interfaces, but distinct patterns emerged in each. Synthetic interfaces exhibited higher interchain/intrachain epistasis importance ratios during early selection rounds, suggesting that interchain interactions played a stronger role in determining binding modes at the onset of selection compared with at the affinity maturation stage (Fig. 5B). These interchain interactions may have provided an initial structural

foothold for binding, enabling intrachain refinements in later rounds to optimize stability and specificity. We speculate that the original high-affinity Z-domain-affibody complex, trapped in a deep energy well that constrains its ability to adapt to the epistasis through reorientation of docking modes, distributes the epistasis more evenly for interface repacking. As a result, natural interfaces maintain a balanced interchain/intrachain epistasis importance ratio throughout selection.

It is instructive to ask how our results on a model system inform our understanding of the initiation step of complex formation during natural protein-protein coevolution. Presumably, naïve proteins acquired mutual affinity through evolution of mutations on complementary surfaces under a selective pressure over a long time period (53). In such a setting, natural coevolution is iterative over time, where advantageous mutations that stabilize the complex accumulate and are fixed, followed by further sampling of mutants until a functional threshold, such as affinity, is met, at which point further affinity-enhancing mutations become superfluous (29, 54, 55). In our system, seed interactions, such as F45^A-F33^B in A7B3 and V14^A-V47^B in A1B1, appear to act as stabilizing anchors, linking seed sequences to refined binding interfaces while preserving evolutionary flexibility, as evidenced by the divergent coevolutionary trajectories. We suggest that such mechanisms likely model mechanisms of protein-protein evolution that occur in nature. During natural protein-protein evolution, it is unlikely that a suitable constellation of mutations would arise simultaneously at the seed stage that would confer high affinity in a single step (56). Thus, we suggest that our model system has captured the essence of a natural mechanism for the initiation of protein-protein complexes.

There are examples of artificially designed protein-protein interfaces that leverage molecular docking followed by side chain design calculations to create new interfaces (57–59). These methods sample shape complementarity extensively through docking but generally lack the ability to capture nuanced side chain-guided interfacial plasticity because of the hierarchical nature of the design procedures as well as challenges associated with overcoming local energy minima. Epistatic clusters of residues are secondary considerations to these types of design methods, typically relying upon the design algorithm to capture them. We used docked poses without specific sequences to set up the residue positions that undergo sequence drifts and came to sequences that convey orthogonality through new binding orientations not easily designable by conventional methods (60, 61).

Artificial intelligence (AI) is making rapid strides in prediction and design of binding proteins (62–65), but these *in silico* approaches remain largely agnostic to the biophysical chemistry and evolutionary mechanisms by which PPIs are formed. Generative AI approaches are effective at producing binders to proteins but are generally most effective at targeting natural binding surfaces that have evolved to bind ligands (59). Both targeting naïve surfaces of proteins and achieving fine specificity among highly similar surfaces remain challenging. Understanding the structural principles underlying protein interfaces at a systems level, using large datasets of paired coevolved protein-protein complexes, could help to obtain deeper insight into molecular recognition and bridge the gaps in current AI-based protein design approaches. Such challenges are difficult for existing models trained largely on static and monomeric structures, which do not explicitly capture how sequence changes reshape interfaces to mediate specificity and cross-reactivity. Our approach might provide insights to construct improved ML-based protein engineering strategies for “undruggable” therapeutic targets (41, 66–68).

Materials and methods

Z domain–Z domain docking

A Z-domain was isolated from the crystal structure of Protein Z in complex with an *in vitro* selected affibody (PDB: 1LP1). The Z-domain structure was minimized using the Rosetta FastRelax protocol, applying

backbone coordinate constraints to preserve the overall fold. A poly-Val variant of the Z-domain, where all amino acids were substituted with valine, was generated using RosettaRemodel. This poly-Val Z-domain was then docked against a second poly-Val Z-domain using PatchDock. Docked models were ranked based on their approximate interface area. Models with an interface area greater than 1000Å² were subjected to refinement with the Rosetta FastRelax protocol under the same constraints. The refined models were ranked by Rosetta Energy Units (REU), and the 25 lowest energy structures were manually inspected in Pymol. The final model was selected based on both energy and interaction mode preference (helices13 interacting with helices23) from the PatchDock set with interface areas exceeding 1000Å².

The original protein sequence was restored onto each Z-domain in the selected docked model, which underwent further refinement with Rosetta FastRelax. Target residues on the interface of one Z-domain were identified based off of this refined model. These target residues were used to generate a rotamer interaction field (RIF) with RIFDock. A second round of PatchDock was performed, targeting the same residues used during the RIF generation step. PatchDock models with an interface area greater than 1000 Å² were used as seeds for RIFDock runs. RIFDock-generated models underwent additional refinement using the FastRelax protocol with coordinate constraints and were ranked by energy. The lowest energy models were visually inspected in Pymol, revealing a convergence to a docked conformation within 0.45Å of the best PatchDock model. This converged model was selected for further sequence design.

The docked model underwent refinement using RosettaDock local-refinement as well as side-chain packing and interface redesign with Rosetta FastDesign. Models were filtered by ddg (threshold = -15) and SASA (threshold = 800). Redesigned models were ranked by energy, and top-ranking designs were analyzed to identify trends in residue preferences, which guided the selection of positions for constructing a degenerate codon library.

Protein expression and purification

The protein Z domain-encoding DNA plasmids were inserted into the pET28 bacterial expression vector. The vector includes the Z-domain gene with either a C-terminal His₆-tag or biotin-acceptor peptide tag (BAP tag, GLNDIFEAQKIEW) followed by His₆-tag, inserted between the *NcoI* and *XhoI* sites of pET28b vector (Novagen). For expression, the plasmids were transformed into *E. coli* BL21 (DE3) cells, which were cultured in TB medium containing 50 mg/L kanamycin at 37°C. When the optical density at 600 nm (OD₆₀₀) reached 0.6, protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG), and the cultures were incubated overnight at 25°C before cell harvesting. Protein purification was carried out using Ni²⁺-NTA agarose column chromatography (Ni-NTA, Qiagen), followed by further purification with size-exclusion chromatography on a Superdex S75 10/300GL Increase column (GE Healthcare). The final protein products were stored in HEPES-buffered saline (HBS; 20 mM HEPES, pH 7.5, 150 mM sodium chloride).

Yeast display of single-chain Z domain dimers

Single-chain Z protein dimers were expressed on the surface of *Saccharomyces cerevisiae* strain EBY100 (Invitrogen, cat. no. C839-00) through fusion to the C terminus of the Aga2 protein. The dimers, linked via a GS-linker containing a 3C protease cleavage site, were positioned between an N-terminal cMyc epitope and a C-terminal HA tag. The construct, formatted as N-cMyc-ZA-linker-ZB-HA-C, was cloned into the pCT302 vector (Addgene #41845). Competent yeast cells were transformed with Z-domain plasmids using a yeast transformation kit (Zymo Research T2001) and plated on SDCAA agar plates (Teknova). Plates were incubated at 30°C for two days until colonies formed. Single colonies were picked and cultured in SDCAA media (pH 4.5, 20 g dextrose, 6.7 g yeast nitrogen base, 5 g bactocasamino acids, 10.4 g sodium citrate and 6.4 g citric acid monohydrate dissolved in 1 liter of

deionized H₂O, supplemented with 10 ml of Gibco™ Penicillin-Sterptomycin, 10,000 U/ml) until reaching an OD₆₀₀ of 10. Cultures were then induced at 20°C for 24 hours by diluting to an OD₆₀₀ of 1.0 in SGCAA medium (similar to SDCAA but containing 20 g galactose instead of dextrose). Protein display levels were validated by staining cells with Alexa Fluor 647-labeled anti-HA antibody (1:50 dilution; Cell Signaling Technology, cat. no. 3444S). Fluorescence signals were analyzed using flow cytometry (Beckman Coulter, CytoFLEX).

On-yeast cleavage-capture assay

For the single-clone cleavage-capture assay, colonies were selected from transformed EBY100 cells that were plated on SDCAA agar plates. A total of 5 × 10⁵ induced yeast cells were stained with Alexa Fluor 647-labeled anti-HA antibody at a 1:50 dilution. After staining, the cells were washed thoroughly with MACS buffer (autoMACS® Running Buffer, Miltenyi, cat. no. 130-091-221) to remove unbound antibodies. The washed cells were then incubated in 20 μL of 3C protease cleavage solution, prepared by diluting lab-made 3C protease to 0.4 mg/mL in MACS buffer, and maintained at 4°C. At each time point, 2 μL of the sample was taken and diluted in 100 μL of ice-cold MACS buffer. The fluorescence intensity of the samples was measured using flow cytometry. To evaluate the affinity between two interacting proteins, the mean fluorescence intensity (MFI) at each time point was normalized to the MFI before cleavage, expressed as a percentage of the maximum MFI.

Yeast displayed libraries

A site-directed mutagenesis library was generated through assembly polymerase chain reaction (PCR) using DNA sequences containing degenerate codons. The gel-purified PCR product was combined with a linearized pCT302 vector and introduced into EBY100 cells via electroporation (69). Following electroporation, the cells were incubated in YPD medium at 30°C for one hour before being transferred to SDCAA media. To assess transformation efficiency, serial dilutions of the recovered cells were plated onto SDCAA agar plates. After a 2-day incubation, protein expression was induced in SGCAA media as described above.

Library DNA sequences are listed below:

GTTGATAATAAATTTAATGCADTSCAATGGDTS GCATTTDTS^{DT}
 SATTTTGCATCTGCCAATTTGAACGAGGAACAGAGAAACGCTTTCA
 TACAGTCTCTAAAAGATGATCCAAGTCAATCAGCAAATTTADTSGC
 CGAAGCTGCGGCCTTAAATGCCGCTCAAGCGCCTAAGGAATTCGGCGG
 AGGTGGGAGSCTGGAAGTTCTGTTCCAGGGTCCGGGAGGCGGCGG
 GAGCGGATCCGTTGACAACAAGTTAACAAGAGCAGCAAATTCGCTT
 TTACGAGATATTACATCTTCCGAATCTTAATGAGATACAGAGGAATDT
 SDTSATTCAGDTSCTGAAAGATGACCCTAGCCAGAGCGCCDTS^{DT}
 CTGGCTDTS^{DT}SGCGAAGATCGCAAACGATGCACAAGCACCTAAA

Theoretical nucleotide diversity: 3.63 × 10⁸

Functional library size: 2.50 × 10⁹

Selection of yeast-display libraries

The yeast-display library was enriched for interacting pairs through a combination of magnetic-activated cell sorting (MACS) and fluorescence-activated cell sorting (FACS), as previously described (27). Initial negative selection was performed using 10 times the theoretical diversity of the library to remove uncleavable variants caused by linker mutations. Positive and negative selections were alternated to enrich interacting pairs while minimizing the accumulation of uncleavable mutants. Library selection involved four rounds of positive selection using MACS (R1 to R4) followed by one round of FACS sorting (R5) to achieve higher purity of interacting pairs. The specific methods for MACS and FACS selections were performed as described previously (27).

Deep sequencing of yeast libraries

DNA was extracted from 5 × 10⁷ to 10 × 10⁷ yeast cells per selection round using the Zymoprep II kit (Zymo Research). Specific six-nucleotide

oligomer (6-mer) barcodes and random 8-mer sequences were incorporated into the flanking regions of the sequencing product through 30 cycles of PCR amplification. The amplified region covered library positions for both Z-A and Z-B. A second PCR step was performed to add Illumina primer sequences, resulting in final products containing the format: Illumina P5-barcode-N8-read-Illumina P7. The final PCR products were purified using agarose gel electrophoresis, quantified with a Nanodrop, and subjected to deep sequencing on an Illumina MiSeq platform with a 2×300 V3 kit.

The amplicons were amplified using the following primers:

Illumina forward primer:

5'-AATGATACGGCGACCACCGAGATCTCACTCTTTCCCTACACGACGCTCTTCCGA-3'

Illumina reverse primer:

5'-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTC-3'

Sequence library filter

To identify enriched oligopeptide pairs, a one-sided hypergeometric test was performed using the *scipy.stats* library. Counts of individual oligopeptides (A and B) and their pairwise combinations (AB) were extracted from the sequencing data. The total population size (M), counts of each A or B oligopeptide (n and N), and observed AB pair counts (k) were used as input parameters for the hypergeometric distribution. The survival function $\{P[X > (k - 1)]\}$ was computed to determine *P* values for each pair. Pairs with *P* values less than 0.05 and counts greater than or equal to 20 were retained as enriched sequences in the round 5 NGS data. The filtered results were stored for further analysis, ensuring that only statistically significant and highly abundant pairs were included in downstream analyses, such as sequence similarity networks and Circos plots.

Sequence similarity network and Circos Plot

Coevolution data were imported from a CSV file containing filtered round 5 NGS data, where *P* values were < 0.05 and read counts were ≥ 20 , resulting in 862 enriched sequences out of a total of 416,872 distinct round 5 sequences, which were used for plotting sequence similarity networks and a Circos plot. Sequence similarity networks and community maps were constructed using the *igraph* software package (70). Nodes in the edit distance-based networks represented specific Z-A/Z-B pairs, and edges were added between nodes when the edit distance between two pairs fell below a predefined threshold.

A Circos plot was generated using the *PyCirclize* library to visualize coevolutionary interactions among protein sequences. A constant score value was assigned for uniform visualization, and interaction categories were color-coded (e.g., red, orange, yellow) to differentiate cluster groups. The final Circos plot provided an effective visualization of complex coevolutionary networks, highlighting sequence clusters and interaction patterns.

X-ray crystallography

Some protein complexes were reductively methylated (71), as noted below, and all were digested with carboxypeptidases A and B, purified by size-exclusion chromatography, and concentrated. All crystals were grown by vapor diffusion using either 100 nl protein with 100 nl well solution or 100 nl protein with 80 nl well solution and 20 nl microseeds, and flash cooled with liquid nitrogen after cryoprotection as noted below.

Methylated A4B1 was concentrated to 245 mg/ml and crystallized against 0.1 M bis-tris-propane pH 9.0 and 30% polyethylene glycol (PEG) 6000. Crystals were cryoprotected by drawing through a drop of paratone N. Diffraction images were collected at the Advanced Light Source (ALS) beamline 8.2.1.

Initial crystals of A2B2 grew at 193 mg/ml from 0.1 M bis-tris pH 5.5 and 25% PEG 3350. These were used to make microseed stock to grow

the final crystals at 228 mg/ml from 0.1 M bis-tris-propane pH 9 and 30% PEG 6000. Crystals were cryoprotected with addition of 30% glycerol. Diffraction images were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 12-1.

Initial crystals of A3B3 were grown at 200 mg/ml from 0.1 M bis-tris pH 6.5 and 25% PEG 3350. A microseed stock from these crystals was used to grow the final crystals from 0.2 M NaCl, 0.1 M bis-tris pH 6.5, and 25% PEG 3350. Crystals were cryoprotected by drawing through a drop of paratone N, and diffraction images were collected at ALS beamline 8.2.1.

Crystals of A5B5 were grown with two cycles of microseeding. Initial seed crystals grew at 210 mg/ml from 0.2 M $\text{Li}_2(\text{SO}_4)$, 0.1 M Tris pH 8.5, and 30% PEG 4000. Secondary seeds were grown with microseeding from 0.1 M HEPES pH 7.5 and 20% PEG 10,000. Final crystals were grown from 0.15 M DL-Malic acid pH 7.0 and 20% PEG 3350. Crystals were cryoprotected by drawing through a drop of paratone N, and diffraction images were collected at ALS beamline 5.0.1.

Methylated A6B6 was concentrated to 168 mg/ml and crystallized from 0.01 M CoCl_2 , 0.1 M MES pH 6.5, and 1.8 M $(\text{NH}_4)_2\text{SO}_4$. Crystals were cryoprotected by drawing through a drop of paratone N, and diffraction images were collected at ALS beamline 2.0.1.

Initial crystals of A7B7 grown at 138 mg/ml from 0.1 M sodium cacodylate pH 6.5 and 1.4 M sodium acetate were used to prepare a microseed stock. Final crystals were grown at 192 mg/ml with microseeding from 0.1 M bis-tris pH 6.1 and 2.1 M $(\text{NH}_4)_2\text{SO}_4$. Crystals were cryoprotected with addition of 3 M sodium malonate pH 6.0, and diffraction images were collected at the National Synchrotron Light Source (NSLS-2) AMX beamline.

Initial seed crystals of methylated A7B3 were grown at 167 mg/ml from 0.1 M Tris pH 8.5 and 25% PEG 3350. Final crystals were grown with microseeding at 136 mg/ml from 0.05 M MgCl_2 , 0.1 M HEPES pH 7.5, and 30% PEG monomethyl ether 550. Crystals were harvested without further cryoprotection and diffraction images were collected at SSRL beamline 12-2.

Diffraction data were indexed, integrated, and scaled using either XDS (72) (A4B1, A2B2, A3B3, A5B5, A6B6, A7B3) or autoproc (73) (A7B7). Space groups were assigned using pointless and reflections were merged with aimless from the CCP4 suite (74, 75). All structures were solved by molecular replacement in Phaser (76) by searching for individual subunits with the following search models: AlphaFold2-generated models of the chains (A3B3, A5B5) (62), side-chain truncated models of the chains from PDB entry 8DA3 (A6B6), side-chain truncated models of the A6B6 chains (A4B1, A2B2, A7B7), and side-chain truncated models of the chains from PDB entry 8DAB (A7B3). Some structures were then rebuilt with phenix.autobuild (77) (A3B3, A6B6, A7B7, A7B3). Additional refinement was performed interactively in Coot (78) and in Phenix (79–82) (A2B2, A3B3, A5B5, A6B6, A7B7) or Buster (83, 84) with final refinement in Phenix (A4B1, A7B3). Final refinement included NCS restraints (A4B1) TLS parameters determined automatically in Phenix (A4B1, A3B3, A5B5, A6B6, A7B7, and A7B3). Model geometry was assessed with Molprobit (85) and contacting residues were identified with Pymol (Schrodinger, LLC) using a distance cutoff of 3.8 Å. Crystallographic software used in this project was compiled and maintained by SBGrid (86). Crystallographic data collection and refinement statistics along with PDB deposition codes are reported in table S1.

Selection of representative cluster structures

To obtain structural representatives for each cluster, we initially attempted to crystallize all seven cluster structures along the diagonal of the specificity matrix in Fig. 4A, along with two high-affinity cross-reactive complexes (A4B1 and A7B3). Among these, the crystal structures of A3B3, A5B5, A6B6, A4B1, and A7B3 were successfully solved. To expand our structural coverage, we selected additional cluster sequences that differed by a single amino acid substitution (one edit distance) from the original representative pairs. These variants were

tested for binding and crystallization, leading to the successful structure determination of A2B2-L14^{AM} for Cluster 2 and A7B7-L43^{BF} for Cluster 7. For Cluster 1, an initial structure solution by molecular replacement at 1.73 Å resolution showed clear density for protein backbones, but most sidechain density was uninterpretable and the structure could not be refined. Closer inspection of the diffraction images identified satellite reflections consistent with incommensurate modulation of the crystal. In light of this result, we leveraged the A4B1 structure as a representative of cluster 1, as it differs from A1B1 by only a single amino acid substitution located outside the interface (fig. S20). Furthermore, we confirmed that the A4B1 structure aligns well with the backbone conformation of A1B1, validating its use as a pseudo-representative structure for Cluster 1.

Surface plasmon resonance

Dissociation constants (K_d) for Z-A/Z-B dimers were measured by SPR on a BIAcore T100 instrument (GE Healthcare). Biotinylated Z-A chains were immobilized on a streptavidin-coated (SA) sensor chip (Cytiva), with a reference channel containing an unrelated protein. All experiments were performed in HBS-P+ buffer (Cytiva) supplemented with 0.5% bovine serum albumin. Serial threefold dilutions of analytes, starting from 50 μM, were injected at a flow rate of 50 μl/min, with an association phase of 120 s followed by a dissociation phase of 180 to 300 s. After each injection, the surface was regenerated with 0.1 M glycine (pH 2.5). Sensorgrams were analyzed using the BIAcore T100 evaluation software, and binding constants were determined by fitting to a steady-state affinity model.

Selection probabilistic model

We developed a statistical ML generative model to model all of the multi-round read count data simultaneously. The core quantity estimated by the model is the probability that any sequence is selected in a given round; the model is aware that output from one selection round is input to the next round. This core, estimated quantity can be thought of as the energy/fitness landscape for that round, and can be aggregated across rounds to obtain an overall energy/fitness function. We call our approach the selection probabilistic model (SPM). SPM is a generative model in that it models the data-generating process of multi-round selection experiments (depicted in fig. S1A) by describing each round in two steps; (i) an update of the sequence distribution from the previous round, where the update arises from the selection process and (ii) generation/sampling of the observed sequence read counts from the updated sequence distribution in the first step, yielding a multinomial likelihood that can be maximized to estimate model parameters. For a detailed description, refer to supplementary text section A. We define $f_r^{\theta}(x)$ as the inferred fitness for round r , with learned parameters θ , as the logarithm of the estimated selection probability (for each sequence and each round), and $f^{\theta}(x)$ as the global (over all rounds) fitness function. The energy is defined as negative fitness. Having inferred these fitness/energy landscapes, we use these for two down-stream applications: (i) epistasis analysis (supplementary text section C) and (ii) simulation of coevolutionary trajectories based on the fitness landscape (supplementary text section D).

SPM training and evaluation

Using the SPM, we learned a function $f_r^{\theta}(x)$ represented as a neural network, to predict the fitness associated with each selection round r for each sequence x (3,650,782 distinct sequences in total). The optimal settings of the parameters θ were determined by maximizing the logarithm of the multinomial likelihood (described in SI Section B) over the sequences in each selection round. Our neural network, $f^{\theta}(x)$, comprised four fully connected hidden layers, each with a ReLU activation function (fig. S1B). To stabilize the training, we applied regularization by weight decay (1×10^{-5}) with the Adam optimizer (87) and a dropout rate of 0.1. The hyperparameters, consisting of number of the

hidden layers (4), hidden dimension (100), learning rate (0.001), batch size (10,000) and regularization strength, were chosen based on the model that produced the lowest MSE between the predicted counts and the experimental observation in a five-fold cross validation on the final selection round (while always using all previous rounds for training). We refer to supplementary text section B for the details of hyperparameter selection and model validation.

Extraction of epistasis importance from SPM-predicted fitness landscape

We use two different types of epistatic characterizations for each epistatic term: (i) the epistatic effect size and (ii) the epistatic importance. As is made more precise in supplementary text section C, the effect size corresponds to the standard effect size one obtains in a linear additive model comprising all epistatic effects of all orders. For example, one epistatic effect size could be the parameter value in the linear additive model for position 2 being an A and position 5 being an L—a second order effect because it involves two positions. By contrast, the epistatic importance refers to how much changes in amino acids at specified positions can change the fitness. For example, one epistatic importance could be the maximum fitness minus the minimal fitness obtained by choosing the least and most favorable amino acids at position 2 and 5; hence, importance is almost like defining the maximally achievable dynamic range at a set of positions, over all possible amino acids. Practically, we use effect size when we're characterizing epistasis related to a particular variant (sequence), whereas when we seek to characterize epistasis for the entire fitness landscape, then importance is the relevant quantity because it is not anchored on any particular variants.

Briefly, we developed an efficient method to extract epistatic information, that is, the importance of different interaction terms; single site, pairwise, and all higher orders, from the learned sequence-to-fitness relationship. In this method, a linearized version of the learned fitness function is constructed that consists of the effect size of different terms (e.g., single site, pairwise, or higher order). This method is related to the Walsh-Hadamard transform of the fitness function (88, 89). After computing the effect sizes of the linearized fitness function, we then define the epistasis importance of a term as the difference between maximum and minimum effect sizes of all possible configurations of amino acids for that term. This choice is inspired by the partial dependence-based variable importance measure for categorical values introduced in the work by Greenwell *et al.* (90). A detailed description of our epistasis-extraction method can be found in supplementary text section C.

Simulating coevolutionary trajectories on the SPM-predicted fitness landscape

To extract the geometry of our estimated binding fitness landscape, we simulated coevolutionary trajectories through it. We simulated coevolutionary trajectories as paths from random points on the landscape that always increase the fitness. Specifically, each move in a trajectory considers all single-position edit distance mutations that increase the fitness, and selects one such single edit at random. This process iterates until no moves are left that increase the fitness, at which point we have arrived at a terminal state, namely, a local minimum. After simulating 10 million such trajectories, we compute the “accessibility” of each sequence as the fraction of trajectories that ended at that sequence. Accessibility of an energy well is defined as the accessibility of that well's lowest-energy sequence. Such an analysis can provide insight into the ruggedness of landscape (the more local minima, the more rugged it is), as well as how different parts of sequence space differ (or not) in their accessibility. Although we simulated 10 million trajectories, fewer trajectories resulted in similar conclusions, suggesting that 10 million was sufficient. More details are provided in supplementary text section D.

Energy landscape well depth, accessibility (width), and their statistical significance

In each energy landscape plot (such as the ones in Fig. 3F and fig. S8, C and D), each well has a specific variant that defines its energy minimum (shown with a colored dot). The width of a well is drawn proportional to the accessibility of that well, which can be seen by tracing out the energy barrier denoted by black lines to adjacent wells. For ease of visualization, we show only wells with the 20 most accessible sequences. For these 20, we chose a random ordering of the wells, with the exception of the natural interface which always centers the deepest energy well, using a random ordering for the rest. Later when we bootstrap this analysis to obtain statistical significance, each bootstrap data set has its own random ordering, thereby revealing that this ordering is unimportant to the major conclusions about differences in landscape geometry between synthetic and natural interfaces.

Although wells are shown adjacent to each other in these visualizations, namely with each well having one neighboring well to the left, and one to the right (other than the two wells on each end), in reality, each well has an energy barrier to every other well (and in particular to the other 19 wells in the plot). Consequently, we define (and show) the depth of an energy well as the minimal cumulative energy barrier that a coevolutionary trajectory needs to overcome from the representative sequence of this well to the representative sequence of any of the other 19 wells.

We use Dijkstra's algorithm to find the path with the minimal energy barrier (details in supplementary text section D).

Having defined the width (accessibility) and depth, we can now define statistical tests related to these quantities in order to gauge if the differences between natural and synthetic interfaces are statistically significant. To do so, for each interface, we sampled 1,000 bootstrap NGS read data sets (each containing the original number of read counts in that data set). For each of these 1,000 data sets, we re-computed our energy landscapes, well accessibility and depths. Specifically, we trained an SPM to predict the fitness landscape—for each of these 1,000 data sets—with which to then simulate 10 million coevolutionary trajectories. We then performed a Mann-Whitney U test for the alternative hypothesis that the depth of the most accessible energy well in the natural interface is larger than 3 times the depth of the deepest energy well among the 20 most accessible wells in the synthetic interface (for 4 times larger, the results were not significant at level $\alpha = 0.05$). Similarly, we performed a Mann-Whitney U test for the alternative hypothesis that the accessibility of the most accessible energy well in the natural interface is larger than 5 times the accessibility of the most accessible energy well in the synthetic interface (for 6 times larger, the results were not significant at level $\alpha = 0.05$).

Identification of seed sequences from coevolutionary trajectories

We require each seed sequence for a specific target sequence to have three properties: (i) Weak but experimentally detectable binding affinity (with SPM-predicted fitness between 18.35 and 22.07 which are the average fitness of sequences with NGS read count of 5 and 20 respectively from the final selection round). (ii) High exclusivity to its strong-binding target sequence, where exclusivity is defined as the ratio of the number of trajectories passing through a given weak-binding sequence and arriving at a specific strong-binding target sequence to the total number of all trajectories passing through the weak-binding sequence (minimum exclusivity = 0.9). (iii) High contribution to its strong-binding target sequence, where contribution is defined as the ratio of the number of trajectories passing through a given weak-binding sequence and arriving at a specific strong-binding target sequence to the total number of all trajectories arriving at the target sequence (minimum contribution = 0.01). Intuitively, sequences with high exclusivity serve as checkpoints in the evolutionary paths to their corresponding target sequences, while sequences with high contribution act as the main origins of the evolutionary paths to their

corresponding target sequences. We focused our analysis on seed sequences that are at least three single-point mutations away from their strong-binding sequences.

Scoring chain pairing specificity and sequence-structure compatibility using Frame2seq

We compute Frame2seq (32) model negative pseudo log likelihoods (PLL) at the interface to assess predicted chain pairing specificity and sequence-structure compatibility. We refer to Frame2seq negative PLL values at the interface as interface scores throughout this work. Given experimentally solved structures and the amino acid sequence at framework positions as input, we introduce a mask at the interface positions and score the clustered library sequences at the masked indices to output interface scores as follows:

$$-\log p\left(x_i = x_i^{\text{library}} \mid x_{-i}^{\text{library}}, S\right) \quad (\text{Eq. 1})$$

where i iterates over interface positions, x_i^{library} is a clustered library sequence, x_{-i} is the sequence x with a mask introduced at i , and S is structure. We predict chain pairing specificity by scoring all Z-A and Z-B chain combinations on corresponding structures. We predict sequence-structure compatibility by scoring all clustered library sequences on all six experimentally solved structures.

Extraction of structure-conditioned epistasis importance using Frame2seq

To extract structure-specific epistasis importance, we utilize the aforementioned epistasis importance extraction method (details described in supplementary text section C) but substitute the fitness function (Eq. 1) $f^0(x)$ from SPM with an approximate structure-conditioned fitness function $f_{F2s}^0(x|S)$ constructed using Frame2seq (where S represents input structural information). We use Frame2seq with published pre-trained weights (32) without fine-tuning. Details of how this function is constructed are provided in supplementary text section E, and our approximations are validated empirically in supplementary text section F.

REFERENCES AND NOTES

1. E. D. Levy, A simple definition of structural regions in proteins and its use in analyzing interface evolution. *J. Mol. Biol.* **403**, 660–670 (2010). doi: [10.1016/j.jmb.2010.09.028](https://doi.org/10.1016/j.jmb.2010.09.028); pmid: [20868694](https://pubmed.ncbi.nlm.nih.gov/20868694/)
2. B. Ma, T. Elkayam, H. Wolfson, R. Nussinov, Protein-protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5772–5777 (2003). doi: [10.1073/pnas.1030237100](https://doi.org/10.1073/pnas.1030237100); pmid: [12730379](https://pubmed.ncbi.nlm.nih.gov/12730379/)
3. P. Gainza et al., Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nat. Methods* **17**, 184–192 (2020). doi: [10.1038/s41592-019-0666-6](https://doi.org/10.1038/s41592-019-0666-6); pmid: [31819266](https://pubmed.ncbi.nlm.nih.gov/31819266/)
4. G. P. Smith, Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985). doi: [10.1126/science.4001944](https://doi.org/10.1126/science.4001944); pmid: [4001944](https://pubmed.ncbi.nlm.nih.gov/4001944/)
5. N. C. Wrighton et al., Small peptides as potent mimetics of the protein hormone erythropoietin. *Science* **273**, 458–464 (1996). doi: [10.1126/science.273.5274.458](https://doi.org/10.1126/science.273.5274.458); pmid: [8662529](https://pubmed.ncbi.nlm.nih.gov/8662529/)
6. W. L. DeLano, M. H. Ultsch, A. M. de Vos, J. A. Wells, Convergent solutions to binding at a protein-protein interface. *Science* **287**, 1279–1283 (2000). doi: [10.1126/science.287.5456.1279](https://doi.org/10.1126/science.287.5456.1279); pmid: [10678837](https://pubmed.ncbi.nlm.nih.gov/10678837/)
7. I. S. Moreira, P. A. Fernandes, M. J. Ramos, Hot spots—A review of the protein-protein interface determinant amino-acid residues. *Proteins* **68**, 803–812 (2007). doi: [10.1002/prot.21396](https://doi.org/10.1002/prot.21396); pmid: [17546660](https://pubmed.ncbi.nlm.nih.gov/17546660/)
8. M. Guharoy, P. Chakrabarti, Conserved residue clusters at protein-protein interfaces and their use in binding site identification. *BMC Bioinformatics* **11**, 286 (2010). doi: [10.1186/1471-2105-11-286](https://doi.org/10.1186/1471-2105-11-286); pmid: [20507585](https://pubmed.ncbi.nlm.nih.gov/20507585/)
9. X. Wang, P. Lupardus, S. L. Laporte, K. C. Garcia, Structural biology of shared cytokine receptors. *Annu. Rev. Immunol.* **27**, 29–60 (2009). doi: [10.1146/annurev.immunol.24.012605.090616](https://doi.org/10.1146/annurev.immunol.24.012605.090616); pmid: [18817510](https://pubmed.ncbi.nlm.nih.gov/18817510/)
10. M. J. Boulanger, A. J. Bankovich, T. Kortemme, D. Baker, K. C. Garcia, Convergent mechanisms for recognition of divergent cytokines by the shared signaling receptor

- gp130. *Mol. Cell* **12**, 577–589 (2003). doi: [10.1016/S1097-2765\(03\)00365-4](https://doi.org/10.1016/S1097-2765(03)00365-4); pmid: [14527405](https://pubmed.ncbi.nlm.nih.gov/14527405/)
11. A. Erijman, E. Rosenthal, J. M. Shifman, How structure defines affinity in protein-protein interactions. *PLoS ONE* **9**, e110085 (2014). doi: [10.1371/journal.pone.0110085](https://doi.org/10.1371/journal.pone.0110085); pmid: [25329579](https://pubmed.ncbi.nlm.nih.gov/25329579/)
 12. J. F. Storz, Compensatory mutations and epistasis for protein function. *Curr. Opin. Struct. Biol.* **50**, 18–25 (2018). doi: [10.1016/j.sbi.2017.10.009](https://doi.org/10.1016/j.sbi.2017.10.009); pmid: [29100081](https://pubmed.ncbi.nlm.nih.gov/29100081/)
 13. C. D. Aakre *et al.*, Evolving new protein-protein interaction specificity through promiscuous intermediates. *Cell* **163**, 594–606 (2015). doi: [10.1016/j.cell.2015.09.055](https://doi.org/10.1016/j.cell.2015.09.055); pmid: [26478181](https://pubmed.ncbi.nlm.nih.gov/26478181/)
 14. Z. Avizemer, C. Martí-Gómez, S. Y. Hoch, D. M. McCandlish, S. J. Fleishman, Evolutionary paths that link orthogonal pairs of binding proteins, Research Square [Preprint] (2023); <https://doi.org/10.21203/rs.3.rs-2836905/v2>
 15. R. Lipsh-Sokolik, S. J. Fleishman, Addressing epistasis in the design of protein function. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2314999121 (2024). doi: [10.1073/pnas.2314999121](https://doi.org/10.1073/pnas.2314999121); pmid: [39133844](https://pubmed.ncbi.nlm.nih.gov/39133844/)
 16. D. de Juan, F. Pazos, A. Valencia, Emerging methods in protein co-evolution. *Nat. Rev. Genet.* **14**, 249–261 (2013). doi: [10.1038/nrg3414](https://doi.org/10.1038/nrg3414); pmid: [23458856](https://pubmed.ncbi.nlm.nih.gov/23458856/)
 17. S. C. Lovell, D. L. Robertson, An integrated view of molecular coevolution in protein-protein interactions. *Mol. Biol. Evol.* **27**, 2567–2575 (2010). doi: [10.1093/molbev/msq144](https://doi.org/10.1093/molbev/msq144); pmid: [20551042](https://pubmed.ncbi.nlm.nih.gov/20551042/)
 18. I. N. Shindyalov, N. A. Kolchanov, C. Sander, Can three-dimensional contacts in protein structures be predicted by analysis of correlated mutations? *Protein Eng.* **7**, 349–358 (1994). doi: [10.1093/protein/7.3.349](https://doi.org/10.1093/protein/7.3.349); pmid: [8177884](https://pubmed.ncbi.nlm.nih.gov/8177884/)
 19. S. D. Dunn, L. M. Wahl, G. B. Gloor, Mutual information without the influence of phylogeny or entropy dramatically improves residue contact prediction. *Bioinformatics* **24**, 333–340 (2008). doi: [10.1093/bioinformatics/btm604](https://doi.org/10.1093/bioinformatics/btm604); pmid: [18057019](https://pubmed.ncbi.nlm.nih.gov/18057019/)
 20. M. Weigt, R. A. White, H. Szurmant, J. A. Hoch, T. Hwa, Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 67–72 (2009). doi: [10.1073/pnas.0805923106](https://doi.org/10.1073/pnas.0805923106); pmid: [19116270](https://pubmed.ncbi.nlm.nih.gov/19116270/)
 21. D. S. Marks *et al.*, Protein 3D structure computed from evolutionary sequence variation. *PLoS ONE* **6**, e28766 (2011). doi: [10.1371/journal.pone.0028766](https://doi.org/10.1371/journal.pone.0028766); pmid: [22163331](https://pubmed.ncbi.nlm.nih.gov/22163331/)
 22. T. A. Hopf *et al.*, Sequence co-evolution gives 3D contacts and structures of protein complexes. *eLife* **3**, 713–724 (2014). doi: [10.7554/eLife.03430](https://doi.org/10.7554/eLife.03430); pmid: [25255213](https://pubmed.ncbi.nlm.nih.gov/25255213/)
 23. D. S. Marks, T. A. Hopf, C. Sander, Protein structure prediction from sequence variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012). doi: [10.1038/nbt.2419](https://doi.org/10.1038/nbt.2419); pmid: [23138306](https://pubmed.ncbi.nlm.nih.gov/23138306/)
 24. D. Ding *et al.*, Co-evolution of interacting proteins through non-contacting and non-specific mutations. *Nat. Ecol. Evol.* **6**, 590–603 (2022). doi: [10.1038/s41559-022-01688-0](https://doi.org/10.1038/s41559-022-01688-0); pmid: [35361892](https://pubmed.ncbi.nlm.nih.gov/35361892/)
 25. A. M. Dean, J. W. Thornton, Mechanistic approaches to the study of evolution: The functional synthesis. *Nat. Rev. Genet.* **8**, 675–688 (2007). doi: [10.1038/nrg2160](https://doi.org/10.1038/nrg2160); pmid: [17703238](https://pubmed.ncbi.nlm.nih.gov/17703238/)
 26. K. B. Levin *et al.*, Following evolutionary paths to protein-protein interactions with high affinity and selectivity. *Nat. Struct. Mol. Biol.* **16**, 1049–1055 (2009). doi: [10.1038/nsmb.1670](https://doi.org/10.1038/nsmb.1670); pmid: [19749752](https://pubmed.ncbi.nlm.nih.gov/19749752/)
 27. A. Yang *et al.*, Deploying synthetic coevolution and machine learning to engineer protein-protein interactions. *Science* **381**, eadh1720 (2023). doi: [10.1126/science.adh1720](https://doi.org/10.1126/science.adh1720); pmid: [37499032](https://pubmed.ncbi.nlm.nih.gov/37499032/)
 28. M. S. Johnson, G. Reddy, M. M. Desai, Epistasis and evolution: Recent advances and an outlook for prediction. *BMC Biol.* **21**, 120 (2023). doi: [10.1186/s12915-023-01585-3](https://doi.org/10.1186/s12915-023-01585-3); pmid: [37226182](https://pubmed.ncbi.nlm.nih.gov/37226182/)
 29. T. N. Starr, J. W. Thornton, Epistasis in protein evolution. *Protein Sci.* **25**, 1204–1218 (2016). doi: [10.1002/pro.2897](https://doi.org/10.1002/pro.2897); pmid: [26833806](https://pubmed.ncbi.nlm.nih.gov/26833806/)
 30. C. M. Miton, K. Buda, N. Tokuriki, Epistasis and intramolecular networks in protein evolution. *Curr. Opin. Struct. Biol.* **69**, 160–168 (2021). doi: [10.1016/j.sbi.2021.04.007](https://doi.org/10.1016/j.sbi.2021.04.007); pmid: [34077895](https://pubmed.ncbi.nlm.nih.gov/34077895/)
 31. M. Högbom, M. Eklund, P.-Å. Nygren, P. Nordlund, Structural basis for recognition by an in vitro evolved affibody. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3191–3196 (2003). doi: [10.1073/pnas.0436100100](https://doi.org/10.1073/pnas.0436100100); pmid: [12604795](https://pubmed.ncbi.nlm.nih.gov/12604795/)
 32. D. Akpinaroglu *et al.*, Structure-conditioned masked language models for protein sequence design generalize beyond the native sequence space. *bioRxiv* 2023.12.15.571823 [Preprint] (2023); <https://doi.org/10.1101/2023.12.15.571823>
 33. D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson, PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **33** (Web Server), W363–W367 (2005). doi: [10.1093/nar/gki481](https://doi.org/10.1093/nar/gki481); pmid: [15980490](https://pubmed.ncbi.nlm.nih.gov/15980490/)
 34. P. S. Huang *et al.*, RosettaRemodel: A generalized framework for flexible backbone protein design. *PLoS ONE* **6**, e24109 (2011). doi: [10.1371/journal.pone.0024109](https://doi.org/10.1371/journal.pone.0024109); pmid: [21909381](https://pubmed.ncbi.nlm.nih.gov/21909381/)
 35. L. A. Clark, H. W. T. van Vlijmen, A knowledge-based forcefield for protein-protein interface design. *Proteins* **70**, 1540–1550 (2008). doi: [10.1002/prot.21694](https://doi.org/10.1002/prot.21694); pmid: [17910054](https://pubmed.ncbi.nlm.nih.gov/17910054/)
 36. J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, Sequence requirements for coiled-coils: Analysis with λ repressor-CN4 leucine zipper fusions. *Science* **250**, 1400–1403 (1990). doi: [10.1126/science.2147779](https://doi.org/10.1126/science.2147779); pmid: [1947779](https://pubmed.ncbi.nlm.nih.gov/1947779/)
 37. B.-Y. Zhu, N. E. Zhou, C. M. Kay, R. S. Hodges, Packing and hydrophobicity effects on protein folding and stability: Effects of β -branched amino acids, valine and isoleucine, on the formation and stability of two-stranded α -helical coiled coils/leucine zippers. *Protein Sci.* **2**, 383–394 (1993). doi: [10.1002/pro.5560020310](https://doi.org/10.1002/pro.5560020310); pmid: [8453376](https://pubmed.ncbi.nlm.nih.gov/8453376/)
 38. W. A. Lim, R. T. Sauer, Alternative packing arrangements in the hydrophobic core of λ repressor. *Nature* **339**, 31–36 (1989). doi: [10.1038/339031a0](https://doi.org/10.1038/339031a0); pmid: [2524006](https://pubmed.ncbi.nlm.nih.gov/2524006/)
 39. H. J. Atkinson, J. H. Morris, T. E. Ferrin, P. C. Babbitt, Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS ONE* **4**, e4345 (2009). doi: [10.1371/journal.pone.0004345](https://doi.org/10.1371/journal.pone.0004345); pmid: [19190775](https://pubmed.ncbi.nlm.nih.gov/19190775/)
 40. M. Krzywinski *et al.*, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009). doi: [10.1101/gr.092759.109](https://doi.org/10.1101/gr.092759.109); pmid: [19541911](https://pubmed.ncbi.nlm.nih.gov/19541911/)
 41. L. Lo Conte, C. Chothia, J. Janin, The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198 (1999). doi: [10.1006/jmbi.1998.2439](https://doi.org/10.1006/jmbi.1998.2439); pmid: [9925793](https://pubmed.ncbi.nlm.nih.gov/9925793/)
 42. R. Norel, S. L. Lin, H. J. Wolfson, R. Nussinov, Shape complementarity at protein-protein interfaces. *Biopolymers* **34**, 933–940 (1994). doi: [10.1002/bip.360340711](https://doi.org/10.1002/bip.360340711); pmid: [8054472](https://pubmed.ncbi.nlm.nih.gov/8054472/)
 43. W. Sheffler, D. Baker, RosettaHoles: Rapid assessment of protein core packing for structure prediction, refinement, design, and validation. *Protein Sci.* **18**, 229–239 (2009). doi: [10.1002/pro.8](https://doi.org/10.1002/pro.8); pmid: [19177366](https://pubmed.ncbi.nlm.nih.gov/19177366/)
 44. J. Shirian, O. Sharabi, J. M. Shifman, Cold Spots in Protein Binding. *Trends Biochem. Sci.* **41**, 739–745 (2016). doi: [10.1016/j.tibs.2016.07.002](https://doi.org/10.1016/j.tibs.2016.07.002); pmid: [27477052](https://pubmed.ncbi.nlm.nih.gov/27477052/)
 45. S. N. S. Gurusinge, B. Oppenheimer, J. M. Shifman, Cold spots are universal in protein-protein interactions. *Protein Sci.* **31**, e4435 (2022). doi: [10.1002/pro.4435](https://doi.org/10.1002/pro.4435); pmid: [36173158](https://pubmed.ncbi.nlm.nih.gov/36173158/)
 46. M. Heyne *et al.*, Climbing Up and Down Binding Landscapes through Deep Mutational Scanning of Three Homologous Protein-Protein Complexes. *J. Am. Chem. Soc.* **143**, 17261–17275 (2021). doi: [10.1021/jacs.1c08707](https://doi.org/10.1021/jacs.1c08707); pmid: [34609866](https://pubmed.ncbi.nlm.nih.gov/34609866/)
 47. R. M. Levy, A. Haldane, W. F. Flynn, Potts Hamiltonian models of protein co-variation, free energy landscapes, and evolutionary fitness. *Curr. Opin. Struct. Biol.* **43**, 55–62 (2017). doi: [10.1016/j.sbi.2016.11.004](https://doi.org/10.1016/j.sbi.2016.11.004); pmid: [27870991](https://pubmed.ncbi.nlm.nih.gov/27870991/)
 48. S. Dutta *et al.*, Determinants of BH3 binding specificity for Mcl-1 versus Bcl-xL. *J. Mol. Biol.* **398**, 747–762 (2010). doi: [10.1016/j.jmb.2010.03.058](https://doi.org/10.1016/j.jmb.2010.03.058); pmid: [20363230](https://pubmed.ncbi.nlm.nih.gov/20363230/)
 49. K. A. Barlow *et al.*, Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein-Protein Binding Affinity upon Mutation. *J. Phys. Chem. B* **122**, 5389–5399 (2018). doi: [10.1021/acs.jpcc.7b11367](https://doi.org/10.1021/acs.jpcc.7b11367); pmid: [29401388](https://pubmed.ncbi.nlm.nih.gov/29401388/)
 50. C. D. Thanos, W. L. DeLano, J. A. Wells, Hot-spot mimicry of a cytokine receptor by a small molecule. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15422–15427 (2006). doi: [10.1073/pnas.0607058103](https://doi.org/10.1073/pnas.0607058103); pmid: [17032757](https://pubmed.ncbi.nlm.nih.gov/17032757/)
 51. K. Li *et al.*, De novo mapping of α -helix recognition sites on protein surfaces using unbiased libraries. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2210435119 (2022). doi: [10.1073/pnas.2210435119](https://doi.org/10.1073/pnas.2210435119); pmid: [36534810](https://pubmed.ncbi.nlm.nih.gov/36534810/)
 52. A. I. Podgoraia, M. T. Laub, Protein evolution. Pervasive degeneracy and epistasis in a protein-protein interface. *Science* **347**, 673–677 (2015). doi: [10.1126/science.1257360](https://doi.org/10.1126/science.1257360); pmid: [25657251](https://pubmed.ncbi.nlm.nih.gov/25657251/)
 53. A. F. Moutinho, F. F. Trancoso, J. Y. Duthel, J. Zhang, The Impact of Protein Architecture on Adaptive Evolution. *Mol. Biol. Evol.* **36**, 2013–2028 (2019). doi: [10.1093/molbev/msz134](https://doi.org/10.1093/molbev/msz134); pmid: [31147689](https://pubmed.ncbi.nlm.nih.gov/31147689/)
 54. E. A. Ortlund, J. T. Bridgman, M. R. Redinbo, J. W. Thornton, Crystal structure of an ancient protein: Evolution by conformational epistasis. *Science* **317**, 1544–1548 (2007). doi: [10.1126/science.1142819](https://doi.org/10.1126/science.1142819); pmid: [17702911](https://pubmed.ncbi.nlm.nih.gov/17702911/)
 55. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114 (2006). doi: [10.1126/science.1123539](https://doi.org/10.1126/science.1123539); pmid: [16601193](https://pubmed.ncbi.nlm.nih.gov/16601193/)
 56. C. Tang, J. Iwahara, G. M. Clore, Visualization of transient encounter complexes in protein-protein association. *Nature* **444**, 383–386 (2006). doi: [10.1038/nature05201](https://doi.org/10.1038/nature05201); pmid: [17051159](https://pubmed.ncbi.nlm.nih.gov/17051159/)
 57. P. S. Huang, J. J. Love, S. L. Mayo, A de novo designed protein-protein interface. *Protein Sci.* **16**, 2770–2774 (2007). doi: [10.1110/ps.073125207](https://doi.org/10.1110/ps.073125207); pmid: [18029425](https://pubmed.ncbi.nlm.nih.gov/18029425/)
 58. S. J. Fleishman *et al.*, Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011). doi: [10.1126/science.1202617](https://doi.org/10.1126/science.1202617); pmid: [21566186](https://pubmed.ncbi.nlm.nih.gov/21566186/)
 59. L. Cao *et al.*, Design of protein-binding proteins from the target structure alone. *Nature* **605**, 551–560 (2022). doi: [10.1038/s41586-022-04654-9](https://doi.org/10.1038/s41586-022-04654-9); pmid: [35332283](https://pubmed.ncbi.nlm.nih.gov/35332283/)
 60. Z. Chen *et al.*, Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111 (2019). doi: [10.1038/s41586-018-0802-y](https://doi.org/10.1038/s41586-018-0802-y); pmid: [30568301](https://pubmed.ncbi.nlm.nih.gov/30568301/)
 61. L. A. Joachimiak, T. Kortemme, B. L. Stoddard, D. Baker, Computational design of a new hydrogen bond network and at least a 300-fold specificity switch at a protein-protein interface. *J. Mol. Biol.* **361**, 195–208 (2006). doi: [10.1016/j.jmb.2006.05.022](https://doi.org/10.1016/j.jmb.2006.05.022); pmid: [16831445](https://pubmed.ncbi.nlm.nih.gov/16831445/)
 62. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). doi: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2); pmid: [34265844](https://pubmed.ncbi.nlm.nih.gov/34265844/)
 63. M. Baek *et al.*, Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021). doi: [10.1126/science.abbj8754](https://doi.org/10.1126/science.abbj8754); pmid: [34282049](https://pubmed.ncbi.nlm.nih.gov/34282049/)

64. J. Dauparas *et al.*, Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **378**, 49–56 (2022). doi: [10.1126/science.add2187](https://doi.org/10.1126/science.add2187); pmid: [36108050](https://pubmed.ncbi.nlm.nih.gov/36108050/)
65. M. Pacesa *et al.*, One-shot design of functional protein binders with BindCraft. *Nature* **646**, 483–492 (2025). doi: [10.1038/s41586-025-09429-6](https://doi.org/10.1038/s41586-025-09429-6).
66. S. Jones, J. M. Thornton, Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13–20 (1996). doi: [10.1073/pnas.93.1.13](https://doi.org/10.1073/pnas.93.1.13); pmid: [8552589](https://pubmed.ncbi.nlm.nih.gov/8552589/)
67. O. Keskin, A. Gursosy, B. Ma, R. Nussinov, Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev.* **108**, 1225–1244 (2008). doi: [10.1021/cr040409x](https://doi.org/10.1021/cr040409x); pmid: [18355092](https://pubmed.ncbi.nlm.nih.gov/18355092/)
68. R. Grünberg, L. Serrano, Strategies for protein synthetic biology. *Nucleic Acids Res.* **38**, 2663–2675 (2010). doi: [10.1093/nar/gkq139](https://doi.org/10.1093/nar/gkq139); pmid: [20385577](https://pubmed.ncbi.nlm.nih.gov/20385577/)
69. G. Chao *et al.*, Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1**, 755–768 (2006). doi: [10.1038/nprot.2006.94](https://doi.org/10.1038/nprot.2006.94); pmid: [17406305](https://pubmed.ncbi.nlm.nih.gov/17406305/)
70. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal*. **1695** (2006).
71. T. S. Walter *et al.*, Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617–1622 (2006). doi: [10.1016/j.str.2006.09.005](https://doi.org/10.1016/j.str.2006.09.005); pmid: [17098187](https://pubmed.ncbi.nlm.nih.gov/17098187/)
72. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010). doi: [10.1107/S0907444909047337](https://doi.org/10.1107/S0907444909047337); pmid: [20124692](https://pubmed.ncbi.nlm.nih.gov/20124692/)
73. C. Vonrhein *et al.*, Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011). doi: [10.1107/S0907444911007773](https://doi.org/10.1107/S0907444911007773); pmid: [21460447](https://pubmed.ncbi.nlm.nih.gov/21460447/)
74. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011). doi: [10.1107/S0907444910045749](https://doi.org/10.1107/S0907444910045749); pmid: [21460441](https://pubmed.ncbi.nlm.nih.gov/21460441/)
75. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013). doi: [10.1107/S0907444913000061](https://doi.org/10.1107/S0907444913000061); pmid: [23793146](https://pubmed.ncbi.nlm.nih.gov/23793146/)
76. A. J. McCoy *et al.*, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007). doi: [10.1107/S0021889807021206](https://doi.org/10.1107/S0021889807021206); pmid: [19461840](https://pubmed.ncbi.nlm.nih.gov/19461840/)
77. T. C. Terwilliger *et al.*, Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 61–69 (2008). doi: [10.1107/S090744490705024X](https://doi.org/10.1107/S090744490705024X); pmid: [18094468](https://pubmed.ncbi.nlm.nih.gov/18094468/)
78. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi: [10.1107/S0907444910007493](https://doi.org/10.1107/S0907444910007493); pmid: [20383002](https://pubmed.ncbi.nlm.nih.gov/20383002/)
79. D. Liebschner *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019). doi: [10.1107/S2059798319011471](https://doi.org/10.1107/S2059798319011471); pmid: [31588918](https://pubmed.ncbi.nlm.nih.gov/31588918/)
80. N. Echols *et al.*, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. D. Adams, Graphical tools for macromolecular crystallography in PHENIX. *J. Appl. Cryst.* **45**, 581–586 (2012). doi: [10.1107/S0021889812017293](https://doi.org/10.1107/S0021889812017293)
81. P. V. Afonine *et al.*, Towards automated crystallographic structure refinement with phenix. *refine. Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012). doi: [10.1107/S0907444912001308](https://doi.org/10.1107/S0907444912001308); pmid: [22505256](https://pubmed.ncbi.nlm.nih.gov/22505256/)
82. J. J. Headd *et al.*, Flexible torsion-angle noncrystallographic symmetry restraints for improved macromolecular structure refinement. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 1346–1356 (2014). doi: [10.1107/S1399004714003277](https://doi.org/10.1107/S1399004714003277); pmid: [24816103](https://pubmed.ncbi.nlm.nih.gov/24816103/)
83. G. Bricogne *et al.*, *BUSTER version 20240710* (United Kingdom Glob; Phasing Ltd, 2017).
84. O. S. Smart *et al.*, Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 368–380 (2012). doi: [10.1107/S0907444911056058](https://doi.org/10.1107/S0907444911056058); pmid: [22505257](https://pubmed.ncbi.nlm.nih.gov/22505257/)
85. V. B. Chen *et al.*, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073); pmid: [20057044](https://pubmed.ncbi.nlm.nih.gov/20057044/)
86. A. Morin *et al.*, Collaboration gets the most out of software. *eLife* **2**, e01456 (2013). doi: [10.7554/eLife.01456](https://doi.org/10.7554/eLife.01456); pmid: [24040512](https://pubmed.ncbi.nlm.nih.gov/24040512/)
87. D. P. Kingma, Adam: A method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG] (2014).
88. A. Aghazadeh, O. Ocal, K. Ramchandran, CRISPR and: Interpretable large-scale inference of DNA repair landscape based on a spectral approach. *Bioinformatics* **36** (suppl. 1), i560–i568 (2020). doi: [10.1093/bioinformatics/btaa505](https://doi.org/10.1093/bioinformatics/btaa505); pmid: [32657417](https://pubmed.ncbi.nlm.nih.gov/32657417/)
89. A. Aghazadeh *et al.*, Epistatic Net allows the sparse spectral regularization of deep neural networks for inferring fitness functions. *Nat. Commun.* **12**, 5225 (2021). doi: [10.1038/s41467-021-25371-3](https://doi.org/10.1038/s41467-021-25371-3); pmid: [34471113](https://pubmed.ncbi.nlm.nih.gov/34471113/)
90. B. M. Greenwell, B. C. Boehmke, A. J. McCarthy, A simple and effective model-based variable importance measure. [arXiv:1805.04755](https://arxiv.org/abs/1805.04755) [stat.ML] (2018).
91. A. Yang, K. C. Garcia, NGS Data for: Structural ontogeny of protein-protein interactions, Dryad (2025). <https://doi.org/10.5061/dryad.79cnp5j8b>.
92. H. Jiang, Codebase and data for structural ontogeny of protein-protein interactions, Zenodo (2025); <https://doi.org/10.5281/zenodo.17282141>.
93. D. Akpınaroglu, T. Kortemme, Structure-conditioned masked language models for protein sequence design generalize beyond the native sequence space (0.0.1), Zenodo (2023). <https://doi.org/10.1101/2023.12.15.571823>.

ACKNOWLEDGMENTS

We thank D. Waghray and A. Velasco for their support and M. Fischbach, J. Xiong, and H. Nisonoff for helpful discussions. We thank M. Allaire at ALS and J. Jakoncic at NSLS-II for assistance with data collection. **Funding:** K.C.G. is an investigator of the Howard Hughes Medical Institute and supported by National Institutes of Health (NIH) grant GM 150125, NIH-R01AI103867, Cancer Grand Challenges partnership financed by CRUK (CGCATF- 2023/100006), the National Cancer Institute (OT2CA297242), the Pew Innovation Fund, and the Yosemite Innovation Fund. J.L. is supported by the Chan Zuckerberg Investigator program. H.J. was supported by the Croucher Fellowship. S.A. was supported by the Swiss National Science Foundation (grant no. P500PT_214430). J.B. was supported by a UC Berkeley Chancellor's Fellowship and National Science Foundation (NSF) Graduate Research Fellowship (DGE 2146752). A.J.L. is supported by an NSF Graduate Research Fellowship and a UCSF Discovery Fellowship. D.A. is supported by an NSF Graduate Research Fellowship. C.P.P. and P.-S.H. are supported by NIH grant R01GM147893. T.K. was supported by NIH grant R35GM145236 and is a Biohub Investigator. D.A. and T.K. benefited from the Microsoft Accelerating Foundation Models Research (AFMR) grant program. The Berkeley Center for Structural Biology is supported in part by the Howard Hughes Medical Institute. The Advanced Light Source is a Department of Energy (DOE) Office of Science User Facility under contract no. DE-AC02-05CH11231. The Pilatus detector on 5.0.1. was funded under NIH grant S10OD021832. The ALS-ENABLE beamlines are supported in part by the NIH, National Institute of General Medical Sciences (NIGMS), grant P30 GM124169. Use of the Stanford Synchrotron Radiation Lightsources, SLAC National Accelerator Laboratory, is supported by the US DOE, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the NIH, NIGMS (P30GM133894). The Center for BioMolecular Structure (CBMS) is primarily supported by the NIH, NIGMS, through grant no. P30GM133893 and by the DOE Office of Biological and Environmental Research FWP no. B0070. This research used resources 17-ID-1 of the National Synchrotron Light Source II, a US DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0012704. This material is based on work supported by the NSF Graduate Research Fellowship Program under grant no. DGE 2146752. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH/NIGMS or the NSF. **Author contributions:** Conceptualization: J.L., K.C.G.; Methodology: A.Y., D.A., S.A., A.J.L., J.B., J.L., K.C.G.; Investigation: A.Y., H.J., D.A., K.M.J., S.A., A.J.L., J.B., C.P.P., L.L., J.L., K.C.G.; Visualization: A.Y., H.J., D.A., K.M.J., S.A., A.J.L., J.B., J.L., K.C.G.; Funding acquisition: T.K., J.L., K.C.G.; Supervision: T.K., J.L., K.C.G.; Writing – original draft: A.Y., H.J., K.M.J., S.A., J.B., J.L., K.C.G.; Writing – review and editing: A.Y., H.J., D.A., K.M.J., S.A., A.J.L., J.B., P.-S.H., T.K., J.L., K.C.G. **Competing interests:** The authors declare that they have no competing interests. **Data, code, and materials availability:** Crystal structures and diffraction intensities have been deposited in the Protein Databank (PDB) with accession IDs [pdb_00009NKM](https://pubmed.ncbi.nlm.nih.gov/00009NKM/), [pdb_00009NKN](https://pubmed.ncbi.nlm.nih.gov/00009NKN/), [pdb_00009NKO](https://pubmed.ncbi.nlm.nih.gov/00009NKO/), [pdb_00009NKP](https://pubmed.ncbi.nlm.nih.gov/00009NKP/), [pdb_00009NKQ](https://pubmed.ncbi.nlm.nih.gov/00009NKQ/), [pdb_00009NKR](https://pubmed.ncbi.nlm.nih.gov/00009NKR/), and [pdb_00009NKS](https://pubmed.ncbi.nlm.nih.gov/00009NKS/). Raw diffraction images for each structure have been deposited in the SGrid databank and are linked to the PDB entries. NGS data are deposited at Dryad (91). The code for training and using SPM, computing epistasis from fitness landscape, and simulating coevolutionary trajectories is available at <https://github.com/hanlunj/coevolution> and is also archived together with training data and model weights at Zenodo (92). The code and model for Frame2seq are available at <https://github.com/dakpinaroglu/Frame2seq> and Zenodo (93). Materials for this study are described in the materials and methods, and reasonable requests can be directed to the corresponding authors. **License information:** Copyright © 2026 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the Author Accepted Manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adx6931
Supplementary Text; Tables S1 and S2; Figures S1 to S21; References (94–101); MDAR Reproducibility Checklist

Submitted 26 March 2025; accepted 16 October 2025

10.1126/science.adx6931



Structural ontogeny of protein-protein interactions

Aerin Yang, Hanlun Jiang, Kevin M. Jude, Deniz Akpinaroglu, Stephan Allenspach, Alex Jie Li, James Bowden, Carla Patricia Perez, Liu Liu, Po-Ssu Huang, Tanja Kortemme, Jennifer Listgarten, and K. Christopher Garcia

Science **391** (6786), eadx6931. DOI: 10.1126/science.adx6931

Editor's summary

A long-standing challenge in protein structure is understanding how protein surfaces enable interactions with other proteins, with the ultimate goal of predicting and designing such surfaces. However, we still do not fully understand how evolution shapes such interactions. Yang *et al.* used synthetic coevolution of artificial protein pairs to explore how protein-protein binding sites form and evolve. They identified seed contacts at a very early stage that could take several different paths as the binder pair coevolved. Natural interface surfaces appear to be preorganized to allow fast evolution of deep energy wells that support binding of new partner proteins. These insights will be important to consider when designing synthetic protein binders and may explain why antibodies and peptides often interact with natural protein binding interfaces. —Michael A. Funk

View the article online

<https://www.science.org/doi/10.1126/science.adx6931>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works