Published in final edited form as:
J Mol Biol. 2007 November 2; 373(4): 941–953.

Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar K_D of a single-domain antibody with a flat paratope

Akiko Koide, **Valentia Tereshko**, **Serdar Uysal**, **Katrina Margalef**, **Anthony A. Kossiakoff**, and **Shohei Koide***
Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637

Abstract

A major architectural class in engineered binding proteins ("antibody mimics") involves the presentation of recognition loops off a single-domain scaffold. This class of binding proteins, both natural and synthetic, has a strong tendency to bind a preformed cleft using a convex binding interface (paratope). To explore their capacity to produce high-affinity interfaces with diverse shape and topography, we examined the interface energetics and explored the affinity limit achievable with a flat paratope. We chose a minimalist paratope limited to two loops found in a natural camelid heavy-chain antibody (V_HH) that binds to ribonuclease A. Ala scanning of the V_HH revealed only three "hot-spot" side chains and additional four residues important for supporting backbone-mediated interactions. The small number of critical residues suggested that this is not an optimized paratope. Using selection from synthetic combinatorial libraries, we enhanced its affinity by >100 fold, resulting in variants with K_d as low as 180 pM with no detectable loss of binding specificity. High-resolution crystal structures revealed that the mutations induced only subtle structural changes but extended the network of interactions. This resulted in an expanded hot-spot region including four additional residues located at the periphery of the paratope with a concomitant loss of the so-called "O-ring" arrangement of energetically inert residues. These results suggest that this class of simple, single-domain scaffolds is capable of generating high-performance binding interfaces with diverse shape. More generally they suggest that highly functional interfaces can be designed without closely mimicking natural interfaces.

Keywords

antibody-antigen interaction; interface topography; binding hot spot; ligand efficiency; O-ring theory

Introduction

Protein-ligand interactions play fundamental roles in biological processes. Consequently, extensive research activities are underway to elucidate the molecular mechanisms underlying protein-ligand interactions and to exploit such knowledge for engineering proteins with novel interaction properties. Minimizing the complexity found in natural protein interaction interfaces is an important and challenging goal in this area of protein engineering. One aspect

*To whom correspondence should be addressed: email, skoide@uchicago.edu; FAX, 1-773-702-0439.

of such minimization is to reduce both the size of the interface and the size of the molecular scaffold (a protein supporting the interface),[1; 2; 3] which complements minimization of chemical complexity in interfaces.[4]

Molecules of this type, also often referred to as antibody mimics, are designed based on the general architecture found in antibodies and other types of receptors where an interaction interface is constructed by multiple surface loops. In minimalist engineering, a single-domain scaffold, either natural or designed, is used for presenting functional residues, instead of larger, more complex natural scaffolds (e.g. the antigen-binding fragment (Fab) of the immunoglobulins). A common architecture of such scaffolds is a β-sandwich that present two or three recognition loops.[1; 2; 3]

Camelids produce a class of antibodies whose antigen recognition unit consists of a monomeric V_H domain, V_HH of about 120 amino acids.[5] Thus, even though it is a natural antibody, a V_HH domain may be viewed as a prototype of single-domain antibody mimics. Because V_HH is monomeric, its binding surface includes only three CDRs.[6] Nevertheless, the V_HH domains commonly show a level of antigen binding affinity similar to that seen for the standard, heterodimeric antibodies that can utilize up to six CDRs. Numerous crystal structures of V_HH in complex with an antigen provide a structural foundation to make V_HH a particularly attractive platform for exploring the capacity of single-domain, β-sandwich scaffolds to generate diverse binding interfaces.

Recent structures of single-domain antibodies and antibody mimics, in complex with their respective targets, show that this class of binding proteins has a strong preference to interact with a preformed cleft in the target protein. A recent, comprehensive study by De Genst et al.[7] showed that anti-hen eggwhite lysozyme (HEL) V_HHs predominantly bind to the active site cleft. Their study found that one anti-HEL V_HH that binds outside the cleft had a comparatively lower affinity (K_d = 70 nM) than the cleft-targeting V_HHs (K_d as low as 77 pM). Likewise, an engineered fibronectin type III domain with a Tyr/Ser-binary interface was found to bind to the cleft of its target, maltose binding protein.[8] This preference to bind to clefts was rationalized in terms of shape complementarity, because a small, single-domain scaffold presents recognition loops in a highly convex shape, which should fit well into a highly concave surface of its target.[7]

The observed bias in the mode of V_HH-antigen interactions suggests that the single-domain antibody mimics might not be suitable for generating a high-affinity interface with a flat geometry. If this were correct, one might wish to use this class of scaffolds exclusively targeting to a concave surface and use/develop other types for targeting a flat surface. As antibody mimics are gaining acceptance as viable antibody alternatives, it is important to define the extent of the potential of this class of scaffold to bind to different topographical features. There are, however, examples of V_HH that bind to a flat or convex epitope,[9; 10] suggesting that single-domain antibodies and antibody mimics are capable of forming diverse interface shapes. Because there has not been detailed energetic characterization of V_HH binding interfaces in general, it is not clear if such flat interfaces are extremely rare members of V_HH paratope repertoire that were particularly successful in optimizing interactions in the paratope (following the convention in antibody-antigen interactions, a paratope denotes the interaction interface of an antibody).

In this work, we performed extensive characterization of a V_HH that has a small, flat interface (Figure 1A) and bind to its antigen, ribonuclease A (RNaseA) with a low nanomolar K_d originally isolated from an immunized animal.[9] This V_HH uses only two CDRs to form a particularly small interface among known V_HHs, thus it can also be considered as a minimalist antibody-like interface in terms of the loop number (an interface consisting of a single

recognition loop probably should be classified as "restrained peptides"). We established a yeast surface display system for V_HH that was used to identify binding hotspots by scanning mutagenesis and to perform affinity maturation leading to variants with ~100 fold higher affinity. A set of high-resolution x-ray crystal structures of the wild-type and affinity matured V_HHs and scanning mutagenesis were used to rationalize the observed affinity enhancement. Together, our results indicate that a single-domain β-sandwich scaffold is capable of producing a high-performance recognition interface with a flat or concave geometry and revealed the synergistic contributions of the overall backbone geometry and of the side chains of a small number of residues. Implications in the minimal requirements for a highly functional protein-protein interaction interface and in protein engineering are discussed.

Results

A high-resolution x-ray crystal structure of anti-RNaseA V_HH

To obtain high-resolution x-ray crystal structures, we constructed an expression system using a synthetic gene for an anti-RNaseA V_HH, cAb-RN05 V_HH [9] (Supplementary Figure 1) that eliminated flexible residues and a His_6-tag at the C-terminus that had been present in the original x-ray crystal structure (1BZQ). For brevity, we will refer to cAb-RN05 V_HH simply as the wild-type V_HH hereafter. Using this system, we determined the x-ray crystal structure of V_HH-RNaseA complex at a 1.4 Å resolution, which revealed a much higher level of details than the original 2.8Å-resolution structure. As expected, the overall structure of the complex is essential identical to the previously determined structure. Additional descriptions of this structure are given near the end of the Results section along with the structure of an affinity-matured variant.

Yeast surface display of V_HH

To facilitate quantitative binding characterization and library sorting, we established a yeast surface display system for the V_HH (Figure 1 and Supplementary Figure 1). We first tested the surface expression of the V_HH and found that the expression level reached the maximum after 18 hours of induction. Unlike yeast display of single-chain Fv (scFv), which commonly requires incubation at lower temperature (20 °C),[11] V_HH can be displayed robustly at 30°C. This observation suggests more efficient folding of the single-domain V_HH molecule than the two-domain scFv, consistent with high-level secretion of a V_HH domain from the yeast.[12] The dissociation constant (K_d) for the antigen, RNaseA, of the V_HH as measured from titration using yeast surface display was 23 ± 3 nM (Figure 1C), consistent with the K_d value of 35nM as previously determined for a purified sample of the same V_HH using surface plasmon resonance by Decanniere et al.[9] These results indicate that the V_HH displayed on the yeast surface is fully functional.

Hot spot identification by alanine-scanning mutagenesis

The thermodynamic contribution of the side chain of each of V_HH residues in the V_HH-RNaseA interface were determined using alanine-scanning mutagenesis.[13] Residues 26-32 in CDR1 and 95-102 in CDR3 were mutated (all numbers are according to the Kabat numbering scheme; Figure 1D). These include all residues that are within 4 Å of RNaseA atoms in the complex and also all the solvent-exposed residues in CDR1 and CDR3. We did not examine CDR2 residues (residues 51-55) because none of them make direct contacts with the antigen. A28 was changed to a Gly, and four Gly residues within the CDRs were respectively replaced with an alanine. All these mutant proteins were expressed on the yeast surface at similar levels. It has been shown that the surface expression level is well correlated with the integrity and stability of a displayed protein.[11] Therefore, these results suggest that the mutations in the CDRs did not disrupt the global fold of the V_HH.

Effects of alanine substitution on antigen binding varied greatly among the mutants (Figure 1D). Critical residues revealed by alanine scanning in CDR1 are Y27, Y31, and I32, all of which make direct contact with RNaseA. G26 has a positive φ angle, but the Ala substitution a minimal effect, suggesting that this position can accommodate a bulkier side chain and possibly changes in backbone conformation.

Among the 12 residues in CDR3, only 4 positions, G95, G96, L99 and Y100d, caused significant increase in K_d (> 10 fold) upon alanine substitution. Among these residues (Figure 1D), only G95 and G96 make a direct contact with the antigen. The Gly-to-Ala mutation increases the size of the side chain and tends to be more disruptive than the other Ala substitutions. In addition, G95 has a positive φ angle, and Ala substitution of this residue may have caused a significant change in the backbone conformation. While it is difficult to compare quantitatively the effects of these Gly-to-Ala mutations with the rest of the mutations tested here, the structure and the mutation results clearly show that the backbone mediated interactions between these Gly residues and RNaseA are important.

The side chains of Y97, R100b and T100c make contacts with RNaseA, but Ala-substitution of these residues resulted in marginal changes. These results are consistent with the structure of the V_HH-RNaseA complex[9] in which backbone atoms of CDR3 residues are involved in antigen interactions (Supplementary Table 1). Although Ala scanning cannot directly probe the energetic contributions of backbone atoms, these backbone-mediated interactions are insensitive to perturbations caused by side chain replacement.

The other two CDR3 residues that showed significant increase in K_d upon mutation, L99 and Y100d, are not directly involved in antigen binding. Rather, their side chains point away from RNaseA to the "side" of the V_HH scaffold, and, together with a highly conserved F37, they form a mini-core that sequesters the hydrophobic surface corresponding to the V_L-binding interface of the V_H domain in the conventional, heterodimeric antibody (Figure 1A). Thus, the large effects by the mutations of these residues can be attributed to their roles as "scaffolding" residues that maintain the proper conformation of CDR3 and hence the paratope. The position equivalent to Y100d was found to be important for maintaining the stability of another V_HH.[14]

Ala scanning revealed that only three side chains in direct contact with the antigen are critical and the identity of an additional four positions are important for backbone-mediated interactions. In this work, we define a hot spot as a position that makes direct contact with the antigen and exhibits >10 fold increase in K_d upon Ala substitution. The hot spot residues revealed by Ala-scanning mutagenesis are localized near the center of the paratope (Figure 1E and F), which are surrounded by less important residues. This architecture is reminiscent of the "O-ring" structure seen in other high-affinity protein-protein interaction interfaces.[15; 16]

Affinity maturation leads to picomolar K_d

To determine whether the affinity of this small, flat paratope of V_HH can be significantly increased, we performed affinity maturation experiments. Experiments were done in two stages as outlined in Figure 2A. First, we constructed separate phage display libraries for the CDR1 residues and the CDR3 residues, respectively. In order to sample a large fraction of possible sequences, we designed libraries so that their theoretical sizes are similar to the size attainable by phage display methods (~10^{10}). Two libraries were made for CDR1. In the first library, termed CDR1-NNT, all positions in CDR1 were randomized using the NNT codon (N is an equal mixture of A, T, G and C) that encodes 15 amino acids (Gln, Glu, Lys, Met and Trp are excluded). The theoretical size of this library is 2.7×10^8. In the second CDR1 library, termed CDR1-NNK, all 20 amino acid types were allowed at each of the "cold spots" (G26, A28, Y29 and T30) while keeping the hot-spot residues unchanged (its theoretical size is 10^6). For CDR3,

we kept the hot-spot residues (G95 and G96) and the scaffolding residues (L99 and Y100d) unchanged and diversified the remaining nine residues using the NNT codon as described above. The theoretical library size (4.2×10^9) was comparable to the actual library size. We sorted these libraries separately to enrich functional CDR1 and CDR3 sequences, which were then combined into a single yeast display library. This yeast display library was sorted based on the off rate to enrich high-affinity clones (Figure 2A).[11]

After library sorting, we identified 24 unique clones from 36 randomly chosen clones (Figure 2B). All of the high-affinity clones originated from the CDR1-NNK library. Many of clones that exhibited a level of binding similar to the wild type originated from the CDR1-NNT library, and the hot-spot residues in the wild type (Y27, Y31 and I32) were completely conserved among them (data not shown), consistent with the Ala scanning results. These results indicate that the Trp residue at position 29 plays an important role in enhancing the affinity, which was encoded in the CDR1-NNK library but not in the CDR1-NNT library).

In CDR3, the amino acid sequences exhibit a high level of convergence at positions that were not part of the hot spots in the wild type (97, 100c 101 and 102), suggesting their importance among the affinity matured clones. In contrast, residues 98, 100, 100a and 100b retained high levels of sequence diversity, suggesting that their side chains are not involved in specific interactions.

We determined K_d values of seven clones using yeast surface display, all of which were in the sub-nM range (Figure 2B). The K_d values of the best clone (clone *a*) was 180 (± 18) pM (Figure 2C), corresponding to an improvement of >100 fold relative to the wild type whose K_d determined using yeast display was 23 nM. Using the yeast surface display format, we estimated the k_{off} of the mutant to be greater than 10 hours and that of the wild type on the order of ~20 minutes (data not shown). These values are consistent with the notion that the reduction of K_d was achieved primarily by a reduction of k_{off}, as expected for an outcome from off-rate based selection.

We produced a soluble protein sample of clone *a* that contains all the consensus residues (Figure 2B). The protein was monomeric as tested with size exclusion chromatography (data not shown). We analyzed its binding kinetics using surface plasmon resonance. Its extremely slow k_{off} made it impossible to precisely determine the K_d, but the data were consistent with the mutant's much lower K_d than the wild type (data not shown).

Clone *a* maintained a high level of binding specificity. We immobilized the wild-type and affinity-matured V_HH samples to agarose beads, and tested interactions between *E. coli* lysate and the immobilized V_HH. We found no significant binding of *E. coli* proteins either to the wild type or the affinity-matured V_HH (Figure 2 D and E), while a “sticky” control (human SUMO4) showed interactions with many types of *E. coli* proteins (Figure 2F). Note that this is a highly stringent test for binding specificity, because only weak affinity with a K_d in the sub μM to low μM ranges, i.e. ~10,000 times higher than the K_d of the mutant V_HH for RNaseA, is required for this type of interaction. Thus, despite the ~100 fold increase in affinity, binding specificity was not compromised in the affinity maturation process.

High-resolution x-ray crystal structures of the wild type and affinity-matured V_HHs complexed with RNaseA

To attempt to elucidate the structural basis of the large magnitude of affinity improvement, we determined the x-ray crystal structure of the affinity matured clone *a* at a 1.9 Å resolution. The overall structure of the affinity-matured complex is nearly identical to the wild-type structure, with the RMSD for the Cα atoms for the V_HH and RNaseA between the two structures of 0.38

and 0.49 Å, respectively. There was, however, a small change in the relative orientation between V_HH and RNaseA (Figure 3A).

The affinity maturation process did not significantly change the backbone conformations of CDR1 and CDR3 (Figure 3B). The RMSD for the Cα atoms for all the CDR1 and CDR3 residues between the two structures was 0.34 Å. The side chains of the conserved residues also showed little conformation changes upon affinity maturation (Figure 3B). Likewise, the epitope residues of RNaseA had very similar conformations between the two complexes (Figure 3C). Only the Y76 side chain had clearly different conformations between the two structures. In the affinity-matured complex, it had two conformers, both of which were distinct from its conformation in the wild-type complex. This movement of Y76 exposed K61 of RNaseA with which the indole side chain of V_HH Y29W interacts (We denote a residue mutated in the affinity complex in the format of (original amino acid)-(position)-(new amino acid) such as Y29W).

The affinity-matured interface buries a slightly smaller amount of surface areas than that of the wild-type complex (Table 1), but it has a slightly larger number of atoms that are in close contact (≤ 4 Å) with the antigen than the wild type. In contrast to the V_HH side of the interface, 37% more antigen atoms are in close contact with the V_HH in the affinity-matured complex, suggesting a more efficient paratope. A small increase of the shape complementarity (SC) value (0.78 vs. 0.76; Table 1) is consistent with this view.

The H-bonds in the interface were highly conserved. In the high-resolution wild-type complex there are nine direct H-bonds at the V_HH–RNaseA interface with six main chain atoms participating in them (Supplementary Table 1). Of these six, five are formed by CDR3 main chain carbonyl groups (G95, G96, R100b and T100c), and one by the NH group of I32 in CDR1. All of these H-bonds are preserved in the affinity matured V_HH complex (Supplementary Figure 2 and Supplementary Table 2). The wild-type V_HH side chains are minimally involved in direct H-bonding interactions with RNaseA. Only Y27 in CDR1 forms direct H-bonds with RNaseA via its side chain, which are also preserved in the affinity-matured complex. There are no salt bridges across the interfaces in either complex.

In the wild-type complex we found four first shell water molecules that coordinate to the CDR main chain and mediate the V_HH-RNaseA interactions (Supplementary Table 2). The equivalent water molecules are found in the affinity-matured complex, but three of them formed additional H-bonds to the main chain groups of V_HH (Supplementary Figure 2 and Supplementary Table 2). Together, the crystal structures revealed that the affinity maturation process did not cause significant structural perturbations to the V_HH structure or to the V_HH-RNaseA interactions that existed in the wild-type complex.

Expanded binding hot spots in the affinity matured V_HH

The strong convergence of amino acid sequences in the affinity-matured clones (Figure 2B) suggests that many of the paratope residues contribute significantly to binding. To quantify these contributions, we performed Ala-scanning analysis of clone *a*, whose x-ray crystal structure had been determined. We analyzed 12 residues that showed a high level of sequence convergence in the selection and excluded three residues that were not diversified in the CDR3 library (G95, G96, and Y100d) and four residues that did not show sequence convergence (positions 98, 100, 100a and 100b; Figure 2B).

In CDR1, the side chain contributions of positions 27 and 29 increased (Figure 5). Although the A28P mutation was common and T30 was highly conserved among the selected clones, their side chains did not significantly contribute to binding, as tested using Ala replacement. Thus, the abundance of the A28P mutation and the conservation of T30 are likely to be due to incomplete sampling of sequences in the library. In CDR3, three residues (Y97D, T100cR and

Q102R) significantly contributed to binding (Figure 4), while their counterparts in the wild-type V_HH only marginally contributed (Figure 1). G101 was fully conserved, but, similar to A28P and T30, Ala substitution had little effect. Therefore, the affinity maturation process converted four "cold spot" residues into hot spots. The hot spot residues account for 76 % of the paratope surface of the affinity-matured V_HH. By comparison, only 35 % of the wild-type paratope surface is from the hot spot residues. Interestingly, the expansion of the hot spots in the affinity-matured V_HH abolished the "O-ring" architecture of the wild-type protein where hot spots are surrounded by neutral residues (compare Figure 1D and Figure 4B).

Structural basis for the contribution of new hot spot residues to antigen binding

As described above, the Ala-scanning analysis identified five residues, Y27, Y29W, Y97D, T100cR and Q102R, that show significantly enhanced contributions to binding in the affinity-matured complex. While rationalizing affinity changes based on interactions observed in the time-averaged structure of the V_HH-RNaseA complexes has limitations, there are several specific, well-defined features that probably play a significant role in the enhanced binding.

The indole ring of Y29W in CDR1 introduces extensive interactions with the side chain of RNaseA K61 (Figure 5A). In the wild-type interface, Y76 of RNaseA folds over K61 and thus masks this hydrophobic patch (Figure 3C).

In both complexes, Y27 hydroxyl forms H-bonds with the backbone NH group of RNaseA N62 and with the backbone carbonyl group of RNaseA Q60. The distance of the latter H-bond is significantly shorter in the affinity-matured complex (2.62 Å) than in the wild type (3.45 Å; Supplementary Table 1). Furthermore, Y27 forms a face-to-edge interaction with the indole ring of Y29W on one side and a face-to-face interaction with Y31 on the other (Figure 5A). These enhanced interactions rationalize the greater impact of the Y27A mutation in the affinity-matured complex (3.3 kcal/mol in the affinity matured complex versus 1.7 kcal/mol in the wild type).

The guanidinium group of T100cR in CDR3 interacts with the carboxyl group of RNaseA E111 and also forms a direct H-bond with the carboxyl group of Y97D in CDR3, another hot spot residue (Figure 5B). This H-bonded Glu-Arg-Asp triad sits on the H-bonded triad of Q69, N71 and C110 of RNaseA, strongly suggesting that this network of polar residues cooperatively forms high-affinity interface. By contrast, the wild-type complex does not contain such a network.

The Q102R side chain atoms do not form direct interactions with other residues, and its importance seems to lie in binding water molecules in the interface. Its guanidinium group forms H-bonds with two water molecules that bridge V_HH and RNaseA. These water molecules are a part of a larger network of H-bonded water molecules (Figure 5C). Taken together, the mutations introduced a number of cooperative networks that seem to synergistically contribute to affinity enhancement without increasing the interface size.

Discussion

Energetics of the V_HH-antigen interfaces

This work represents the first set of Ala-scanning analysis of V_HH-antigen interfaces to characterize their binding determinants. As expected from the extensive involvement of the V_HH backbone atoms in antigen binding, Ala-scanning of the wild-type V_HH revealed that a relatively few side chains (Y27, Y31 and I32) make critical direct interactions with the antigen (Figure 1). Among the residues involving in backbone-mediated contacts, only G95 and G96 did not tolerate Ala substitution. By comparison, eight to ten side chains are found to be hot spots (>10 fold increase in K_d upon Ala-substitution) in the paratopes of conventional,

heterodimeric antibodies[17; 18; 19; 20]. Similarly, nine residues are hot spots in the growth hormone receptor that consists of loops from two domains.[15] Thus, the wild-type V_HH paratope achieves high-affinity interaction using much fewer hot spot residues than these examples of larger interfaces.

The high efficiency of the small V_HH interface is manifested in its shape complementarity (Table 1). Its shape compelmentarity (SC) value (0.76) is much higher than 0.64-0.68 found for conventional antibodies and at the high end of the range (0.70-0.76) for oligomeric proteins and protease-inhibitor complexes.[21] Interestingly, we found that other V_HH interfaces also have very high SC values regardless of the paratope shape (Table 1), suggesting that the camelid immune system is capable of generating small and high-efficiency interfaces in V_HH in general. A higher shape complementarity may be a general requirement for achieving high affinity using a smaller paratope. Alternatively, because of the single-domain architecture with fewer CDR loops, amino acid diversity at each position of V_HH may be more extensively sampled in the camelid immune system compared with that in the conventional antibodies where many more positions are diversified. Better sampling may lead to the selection of better-fitted CDR residues.

By extensively sampling CDR sequencing, we were able to enhance the affinity of the V_HH to 150 pM. The affinity enhancement was not achieved by simply expanding the interface area. Rather, the enhancement seems to be the results of the formation of cooperative network among the interface residues that collectively improve various types of biophysical factors including packing, hydrophobic interaction and H-bonding. This view is consistent with the conclusion from comprehensive studies on the mechanism of affinity maturation for the T cell receptor Vβ domain-superantigen interface where packing, hydrophobic burial, bridging water molecules and electrostatic complementarity are all improved by mutations.[22; 23]

As the affinity was improved, four “cold spot” positions became hot spots (Figure 4). The hot spot residues constitute 76% of the paratope surface, and this expansion of hot spot residues completely eliminated the “O-ring” of energetically inert residues. The affinity-matured V_HH retained a very high level of binding specificity, as it did not interact with E. coli proteins (Figure 2). Together, These results indicate that the O-ring architecture is not a prerequisite either for high affinity or high specificity.

Recently, Thanos et al.[24] demonstrated that a small molecule, SP4206, binding to a flat surface of interleukin 2 (IL2) has a much higher “ligand efficiency” (the binding free energy per contact atom)[25] than that of IL2 receptor that binds to the same surface. Their study thus has set a benchmark for two distinct types of protein-ligand interactions. Interestingly, we found that the ligand efficiency of wild-type cAb-RN05 was identical to that of SP4206 (Table 1), indicating that the V_HH is a highly efficient ligand. The affinity-matured V_HH essentially maintained the ligand efficiency, and a large increase in contacting atoms resulted in much higher affinity (Table 1). Furthermore, we found that some of other V_HHs have even higher ligand efficiency values (Table 1). Thus, these results clearly demonstrate that a flat interface built from a single-domain β-sandwich scaffold can be as efficient as drug-like small molecules.

Implications to interface engineering

The characteristics of the distribution of binding energies in the wild-type and affinity matured interfaces provide important insights into factors governing protein-protein interaction interfaces in general and a guideline for engineering a synthetic interface.

The large effects of replacing two “scaffolding” residues in CDR3 (L99 and Y100d) illustrate the importance of the CDR conformation in generating a high-affinity interface. They form hydrophobic interactions with the V_HH scaffold, which are critical in defining the fold of

CDR3. It is likely that these scaffolding residues are also important for the high specificity of the wild-type V_HH. In the wild-type paratope, only three side chains are involved in critical contacts. Consequently, interactions mediated by the V_HH backbone atoms must make large contribution to high-affinity interaction with the antigen and at the same time to prevent binding to unrelated molecules. Our results strongly suggest that L99 and Y100d play the central role in defining the conformation of CDR backbone that is suitable for binding to the RNaseA epitope, but incompatible with other surfaces. This notion in turn suggests that the positions and directions of peptide backbone functional groups can create a large ensemble of high-affinity and -specificity interfaces. This view is consistent with recent examples of "minimalist chemical diversity" interfaces that have demonstrated the importance of conformational diversity over chemical diversity in generating protein-protein interaction interfaces.[4; 8; 26] The position equivalent to Y100d near the C-terminus of CDR3 is usually occupied by an aromatic amino acid among V_HHs, and thus, residues from this position to the C-terminus of CDR3 do not exhibit a large degree of conformational diversity. In contrast, L99 at the center of CDR3 is not conserved. Thus, CDR3 has a capacity to generate diverse conformations by simply moving a scaffolding residue along the loop and also by changing the loop length.

CDR1 and CDR3 of the V_HH studied in this work contribute approximately equally to binding. In the conventional antibodies, CDR-H3 is located near the center of the paratope, it has the greatest sequence diversity among the CDRs because of the V(D)J recombination mechanism, and usually CDR-H3 residues make the largest contribution to binding. The natural V_HH sequence diversity is also produced with the same recombination mechanism[6] and thus, CDR3 has the largest sequence diversity. The hot spot residues of CDR1 and CDR3 of the anti-RNaseA V_HH are located near the center of the paratope, as seen in other interfaces.[27; 28] Thus, there is a mismatch between the interface architecture and the focus of sequence diversity. We predict that synthetic V_HH libraries containing approximately equal distribution of amino acid diversity between CDR1 and CDR3 would outperform the CDR3-focused diversity of the natural repertoire. In other words, the geometry of the interface should dictate the design of synthetic libraries, rather than the idiosyncrasy associated with the heritage of a scaffold.

Our results clearly indicate that a small, flat interface consisting of only two surface loops is capable of producing flat binding interfaces with high affinity, high specificity and high spatial efficiency. The interface has relatively small number of essential positions, and thus engineering this type of interface from combinatorial libraries should not be particularly difficult. With demonstrated ability of this class of molecules to target a preformed cleft, our results strongly suggest the versatility of this class of scaffolds in generating diverse interfaces. These interfaces require the synergy of optimized backbone conformation and optimized placement of contacting side chains. Because both of these parameters are controlled by the amino acid sequences, a successful strategy should maximize the shape diversity and enrich amino acids that are poised to form molecular contacts in the loop sequences. This view is also supported by recent successes of the "restricted amino acid code" strategy for engineering protein-interaction interfaces.[4]

Experimental Procedures

V_HH gene construction and yeast surface display

The synthetic gene for cAb-RN05 V_HH was constructed by PCR using seven synthesized oligonucleotides. The PCR product was digested with BamHI and XhoI, and ligated between the BamHI and XhoI sites of the pGalAga vector that had been constructed by removing the Xpress tag from pYD1 (Invitrogen). We removed the Xpress tag because of its cross-reactivity with anti-FLAG tag antibodies (Sigma). The resulting plasmid, pGalAgaCamR, encodes a fusion protein of Aga2-V_HH-V5 tag-His tag I. The GAL1 promoter regulates the expression

of the fusion protein. Mutants with alanine substitution for alanine scanning were constructed using standard PCR methods.

RNaseA, monoclonal anti-V5 IgG and anti-mouse IgG-fluorescein isothiocyanate (FITC) conjugate were purchased from Sigma. Streptavidin-PE conjugate was purchased from Molecular Probes. RNaseA was biotinylated using EZ-Link NHS-LC-Biotin from Pierce.

The expression level of the V_HH and the binding to RNaseA were measured by flow cytometry using a FACScan instrument (Beckton Dickinson) according to Boder et al.[11] Yeast cells harboring the pGalAgaCamR plasmid or its variant were grown at 30 °C for 18 to 28 hours in media that contains 2% galactose. The cells were incubated with monoclonal anti-V5 antibody, followed by FITC-conjugated anti-mouse antibody. The binding was detected by positive signal of PE after incubating cells with biotinylated RNaseA, followed by PE-conjugated streptavidin.

Library sorting

The phage display vector for V_HH, pCamR-V5p3, was constructed by replacing the gene for the fibronectin type III domain in the phage display vector, pAS38,[29] with the DNA segment encoding V_HH and the V5 epitope tag taken form the V_HH yeast surface display vector described above. Phage particles were prepared in the presence of 0.2mM IPTG using previously described procedures [30].

Combinatorial phage display libraries were made using a high-efficiency Kunkel mutagenesis method.[31] RNaseA was biotinylated with sulfosuccinimidyl 2-(biotinamido)-ethyl-1,3-dithiopropionate (EZ-Link Sulfo-NHS-SS-Biotin; Pierce Biotechnologies). Phage display libraries were sorted as follows. A library solution was incubated with 50 nM of biotinylated RNaseA for 30 minutes, and then the biotinylated RNaseA was captured with streptavidin-coated magnetic beads (Streptavidin MagneSphere Pramagnetic Particles; Promega). After washing the beads three times with TBST (150 mM NaCl, 50 mM Tris-HCl, pH7.5 and 0.5% Tween20), the phages were eluted by cleaving a linker within the biotinylation reagent with 100 mM DTT dissolved in 20 mM Tris-HCl, pH 8.0. Eluted phages were amplified and library sorting was repeated using 25 and 10 nM of RNaseA for the second and third rounds, respectively.

The enriched sequences for CDR1 and CDR3 were then combined using PCR methods, and the resulting ensemble of V_HH genes was inserted into the yeast surface vector using the in vivo recombination method.[11]

Specificity test

The wild-type and affinity matured V_HH proteins each containing an N-terminal His_6 tag (80 μg) were mixed with 250 μl of E. coli lysate corresponding to a cell density of 25 OD_{600nm}, and this mixture was purified in a 1 ml Ni-affinity cartridge (HisTrap column; Amersham). Fractions were analyzed using SDS-PAGE. As a control, human SUMO4 with an N-terminal His_6 tag (a gift of Drs. I. Dementieva and S. Goldstein, the University of Chicago) was treated in the same manner.

Crystallization

RNaseA was dissolved in 20mM TBS at a concentration of 100 mg/ml. The purified V_HH proteins were dialyzed against 20mM TBS, concentrated to ~20mg/ml and mixed with RNaseA protein at a 1:2 ratio to form a complex. The complex was purified with Superdex75 gel filtration column. The wild-type complex was eluted in 20mM Na-phosphate buffer, pH 8.0, 150mM NaCl and concentrated to 20mg/ml using Amicon ultrafiltration devices with 5kDa

cutoff membranes. The affinity-matured V_HH complex was eluted in 10mM Tris-HCl, pH 8.0, 150mM NaCl and concentrated to 15-20mg/ml.

Crystals of V_HH complexes with RNaseA were grown by vapor diffusion in hanging drops at 19°C. The crystals of wild-type complex were obtained from solution containing 0.1 M Tris-HCl, pH 8.5, 0.2 M $MgCl_2$, 25% PEG 3350. The single crystals ($0.3\times0.15\times0.1$ mm^3) formed within 1-2 days and typically grew to the final dimensions in one week. The affinity matured complex was crystallized from solution containing 0.1 M Bis-Tris, pH 5.5, 0.2 M $(NH_4)_2SO_4$, 25% PEG 3350 and produced thin plate clusters. Our attempts in optimizing the crystallization conditions to obtain single crystals were unsuccessful. For x-ray data collection the crystal clusters were broken into several pieces (0.1 mm × 0.1 mm). All crystals were flash-frozen in liquid nitrogen from the mother liquid supplemented with 20% of glycerol.

X-ray diffraction data collection and structure determination

X-ray diffraction data were collected at 100 K using the Advanced Photon Source (Argonne National Laboratory). The x-ray data collection for wild type complex was performed on one single crystal at the BioCARS 14BM beam line (wavelength l=0.900 Å) using 2 degrees (5 *sec* exposure) and 3 degrees (1 *sec* exposure) oscillation width per frame in high and low resolution diffraction experiments, respectively. The x-ray data for the affinity-matured complex were obtained from two crystal pieces at GM/CA 23ID beam line using 0.97936 Å wavelength, 2-second exposure time and 0.5 degrees oscillation width. The X-Ray data were processed with HKL2000.[32] The structures were solved by molecular replacement using MOLREP [33] and x-ray data high-resolution cutoff of 3 Å. The PDB entry 1BZQ [9] was used as a trial model in the rotation and translation search for the wild-type complex. To solve the affinity-matured structure we used the refined wild-type structure. Similar refinement protocols were used for both complexes. R_{free} was monitored by setting aside 5% of the reflection as a test set. Rigid-body, positional refinement and stimulated annealing were performed in CNS1.1.[34] The resolution of x-ray data was increased gradually and all data were included in the last cycles. The models were checked manually against a composite omit map calculated in CNS1.1. The positional and temperature factor refinement was completed with the program REFMAC5 [35] incorporated in CCP4 [36]. The SigmaA-weighted 2Fobs-Fcalc and Fobs-Fcalc Fourier maps were calculated using CCP4. The search for solvent molecules was performed with help of ARP-WARP program.[37] The Fourier maps were displayed and examined in TURBO-FRODO.[38] The crystal data, data collection and refinement statistics are summarized in Table 2. A Ramachandran plot calculated with PROCHECK[39] indicates that 99.5% of the non-Gly and non-Pro residues in both complexes lie in the most favored and additional allowed regions. Figures showing the electron density map and three-dimensional structures were prepared using TURBO-FRODO and RIBBONS.[40] The coordinates and structure factors have been deposited in the Protein Data Bank with an entry code 2P49 and 2P4A for the wild-type and affinity-matured complexes, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported in part by NIH grant R01-GM072688 and by the University of Chicago Cancer Research Center. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38 and DE-AC02-06CH11357. Portions of this work were performed at the BioCARS and GM/CA -CAT located at Sector 14BM and 23ID, respectively, of the Advanced Photon Source. Use of the BioCARS was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707. Use of the GM/CA-CAT was funded with Federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104).

References

1. Binz HK, Amstutz P, Pluckthun A. Engineering novel binding proteins from nonimmunoglobulin domains. Nat Biotechnol 2005;23:1257–68. [PubMed: 16211069]
2. Binz HK, Pluckthun A. Engineered proteins as specific binding reagents. Curr Opin Biotechnol 2005;16:459–69. [PubMed: 16005204]
3. Hosse RJ, Rothe A, Power BE. A new generation of protein display scaffolds for molecular recognition. Protein Sci 2006;15:14–27. [PubMed: 16373474]
4. Sidhu SS, Kossiakoff AA. Exploring and designing protein function with restricted diversity. Curr Opin Chem Biol. 2007
5. Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R. Naturally occurring antibodies devoid of light chains. Nature 1993;363:446–8. [PubMed: 8502296]
6. Muyldermans S, Cambillau C, Wyns L. Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. Trends Biochem Sci 2001;26:230–5. [PubMed: 11295555]
7. De Genst E, Silence K, Decanniere K, Conrath K, Loris R, Kinne J, Muyldermans S, Wyns L. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. Proc Natl Acad Sci U S A 2006;103:4586–91. [PubMed: 16537393]
8. Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S. High-affinity single-domain binding proteins with a binary-code interface. Proc Natl Acad Sci U S A 2007;104:6632–7. [PubMed: 17420456]
9. Decanniere K, Desmyter A, Lauwereys M, Ghahroudi MA, Muyldermans S, Wyns L. A single-domain antibody fragment in complex with RNase A: noncanonical loop structures and nanomolar affinity using two CDR loops. Structure Fold Des 1999;7:361–70. [PubMed: 10196124]
10. Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D, Larsson G, Spencer A, Archer DB, Sasse J, Muyldermans S, Wyns L, Redfield C, Matagne A, Robinson CV, Dobson CM. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. Nature 2003;424:783–8. [PubMed: 12917687]
11. Boder ET, Wittrup KD. Yeast surface display for directed evolution of protein expression, affinity, and stability. Methods Enzymol 2000;328:430–44. [PubMed: 11075358]
12. Frenken LG, van der Linden RH, Hermans PW, Bos JW, Ruuls RC, de Geus B, Verrips CT. Isolation of antigen specific llama VHH antibody fragments and their high level secretion by Saccharomyces cerevisiae. J Biotechnol 2000;78:11–21. [PubMed: 10702907]
13. Wells JA. Systematic mutational analyses of protein-protein interfaces. Methods Enzymol 1991;202:390–411. [PubMed: 1723781]
14. Bond CJ, Marsters JC, Sidhu SS. Contributions of CDR3 to V H H domain stability and the design of monobody scaffolds for naive antibody libraries. J Mol Biol 2003;332:643–55. [PubMed: 12963373]
15. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. Science 1995;267:383–386. [PubMed: 7529940]
16. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol 1998;280:1–9. [PubMed: 9653027]
17. Lang S, Xu J, Stuart F, Thomas RM, Vrijbloed JW, Robinson JA. Analysis of antibody A6 binding to the extracellular interferon gamma receptor alpha-chain by alanine-scanning mutagenesis and random mutagenesis with phage display. Biochemistry 2000;39:15674–85. [PubMed: 11123892]
18. Pons J, Rajpal A, Kirsch JF. Energetic analysis of an antigen/antibody interface: alanine scanning mutagenesis and double mutant cycles on the HyHEL-10/lysozyme interaction. Protein Sci 1999;8:958–68. [PubMed: 10338006]
19. Vajdos FF, Adams CW, Breece TN, Presta LG, de Vos AM, Sidhu SS. Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis. J Mol Biol 2002;320:415–28. [PubMed: 12079396]
20. Li Y, Urrutia M, Smith-Gill SJ, Mariuzza RA. Dissection of binding interactions in the complex between the anti-lysozyme antibody HyHEL-63 and its antigen. Biochemistry 2003;42:11–22. [PubMed: 12515535]

21. Lawrence MC, Colman PM. Shape complementarity at protein/protein interfaces. J Mol Biol 1993;234:946–50. [PubMed: 8263940]
22. Cho S, Swaminathan CP, Yang J, Kerzic MC, Guan R, Kieke MC, Kranz DM, Mariuzza RA, Sundberg EJ. Structural basis of affinity maturation and intramolecular cooperativity in a protein-protein interaction. Structure 2005;13:1775–87. [PubMed: 16338406]
23. Yang J, Swaminathan CP, Huang Y, Guan R, Cho S, Kieke MC, Kranz DM, Mariuzza RA, Sundberg EJ. Dissecting cooperative and additive binding energetics in the affinity maturation pathway of a protein-protein interface. J Biol Chem 2003;278:50412–21. [PubMed: 14514664]
24. Thanos CD, DeLano WL, Wells JA. Hot-spot mimicry of a cytokine receptor by a small molecule. Proc Natl Acad Sci U S A 2006;103:15422–7. [PubMed: 17032757]
25. Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proc Natl Acad Sci U S A 1999;96:9997–10002. [PubMed: 10468550]
26. Fellouse FA, Li B, Compaan DM, Peden AA, Hymowitz SG, Sidhu SS. Molecular recognition by a binary code. J Mol Biol 2005;348:1153–62. [PubMed: 15854651]
27. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285:2177–98. [PubMed: 9925793]
28. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 2002;12:14–20. [PubMed: 11839484]
29. Koide A, Bailey CW, Huang X, Koide S. The fibronectin type III domain as a scaffold for novel binding proteins. J Mol Biol 1998;284:1141–1151. [PubMed: 9837732]
30. Sidhu SS, Lowman HB, Cunningham BC, Wells JA. Phage display for selection of novel binding peptides. Methods in Enzymology 2000;328:333–63. [PubMed: 11075354]
31. Sidhu SS, Lowman HB, Cunningham BC, Wells JA. Phage display for selection of novel binding peptides. Methods Enzymol 2000;328:333–63. [PubMed: 11075354]
32. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–26.
33. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. J Appl Cryst 1997;30:1022–5.
34. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921. [PubMed: 9757107]
35. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–55. [PubMed: 15299926]
36. Collaborative_Computational_Project. The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst 1994;D50:760–763.
37. Zwart PH, Langer GG, Lamzin VS. Modelling bound ligands in protein crystal structures. Acta Crystallogr D Biol Crystallogr 2004;60:2230–9. [PubMed: 15572776]
38. Cambillau, C.; Roussel, A. Turbo Frodo, version OpenGL.1. University Aix-Marseille II; Mrseille: 1997.
39. Laskowski RA, MacArthur MW, Moss D, Thornton JM. PROCHECK: a program to check the sterochemical quality of protein structures. J Appl Cryst 1993;26:283–291.
40. Carson M. Ribbons. Methods Enzymol 1997;277:493–505. [PubMed: 18488321]
41. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004;14:1188–90. [PubMed: 15173120]
42. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 1990;18:6097–100. [PubMed: 2172928]

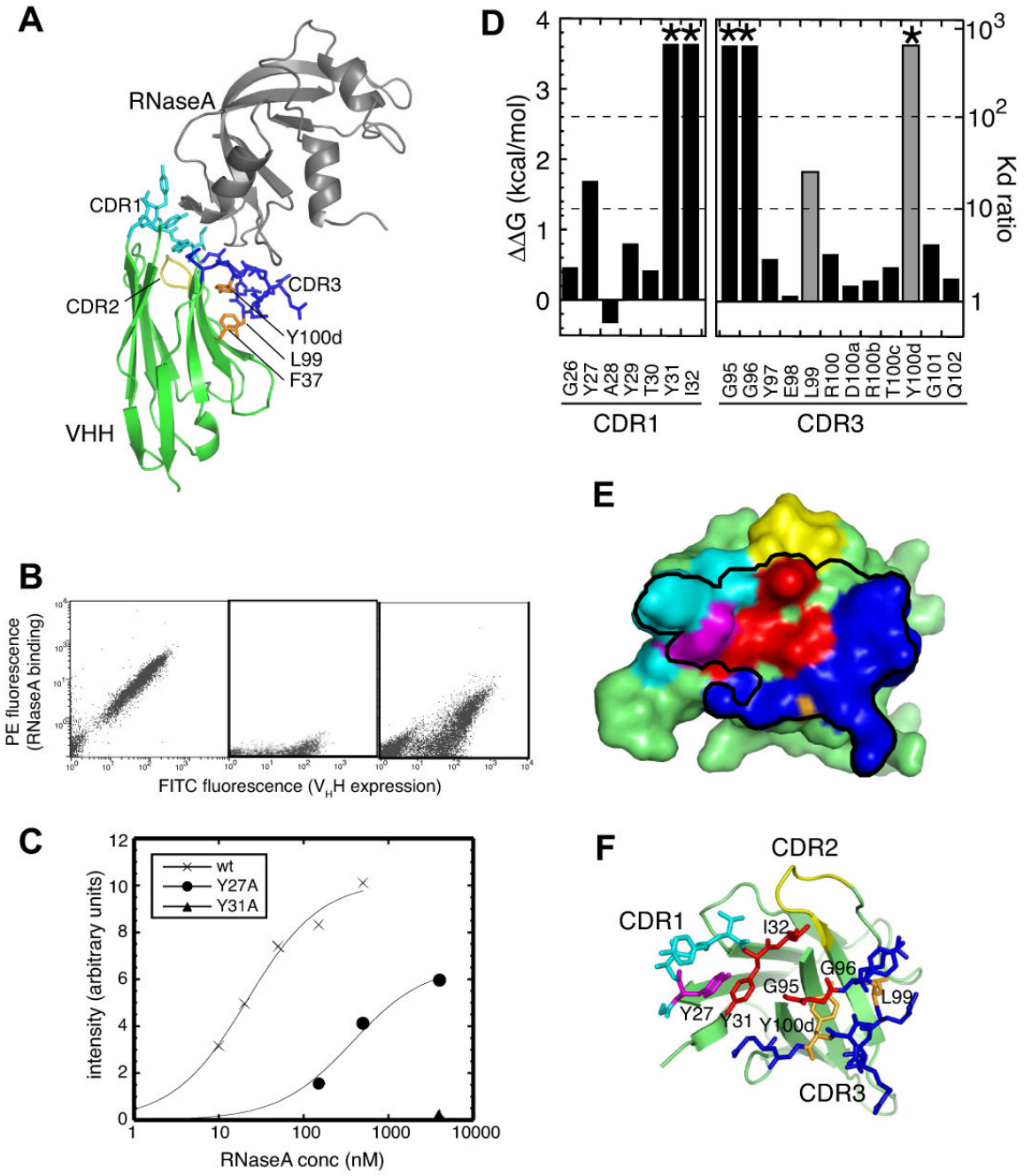

Figure 1. Yeast surface display characterization of the surface expression and antigen binding of cAb-RN05 V_HH and Ala-scanning mutants

(**A**) Ribbon drawing of cAb-RN05 –V_HH-RNaseA complex (PDB ID, 1BZQ). V_HH is shown in green and RNaseA is in gray. CDR1, 2 and 3 of V_HH are shown as stick models and in cyan, yellow and blue, respectively. The side chains of F37, L99 and Y100d, which form a mini core, are shown in orange and labeled.

(**B**) Flow cytometry analysis of yeast cells displaying cAb-RN05 V_HH. The horizontal axis shows the intensity of FITC fluorescence emission that indicates the degree of surface expression of V_HH. The vertical axis shows the intensity of PE fluorescence emission that indicates the degree of antigen (RNaseA) binding. Each dot represents data for a single cell.

The left panels shows data taken with 1 μM RNaseA, the center panel shows the data without the antigen added to the reaction (negative control), and the right panel shows data with 1 μM RNaseA for the Y27A mutant that has reduced antigen binding affinity.

I Antigen titration of the wild-type and mutant V_HHs as measured in the yeast surface display format. The sum of PE fluorescence intensity for 10,000 cells is plotted as a function of RNaseA concentration. The X's show data for wild type V_HH, the circles for the Y27A mutant, and the triangle for the Y31A mutant. Y31A showed essentially no binding. The curves show the best fit of the 1:1 binding model.

(**D**) Effects of Ala substitution on the affinity of V_HH-antigen interaction, as expressed in ΔΔG (labeled on the left axis) and in the ratio of K_d (= $K_d^{\text{wild-type}}/K_d^{\text{mutant}}$; labeled on the right axis). *The K_d for Y31A, I32A, G95A, G96A, and Y100dA is estimated to be greater than 40 μM. The amino acid residue numbers are according to the Kabat convention.[9] Data for residues that do not directly contact the antigen (L99 and Y100d) are shown in gray.

(**E** and **F**) The surface representation of wild-type cAb-RN05 V_HH showing the paratope and binding hot spots. The black line encloses the surface of residues within 5Å from RNaseA. Residues where Ala substitution increased the K_d by >100 fold and by 10–100 fold are show in red and in magenta, respectively. L99 and Y100d, for which Ala substitution significantly reduced antigen binding but which are not in direct contact with the antigen, are shown in orange. The remaining residues in CDR1 and CDR3 are shown in light blue and dark blue, respectively. The CDR2 residues are in yellow. The V_HH scaffold is shown as a cartoon model and CDR1 and CDR3 residues as stick models in (E).

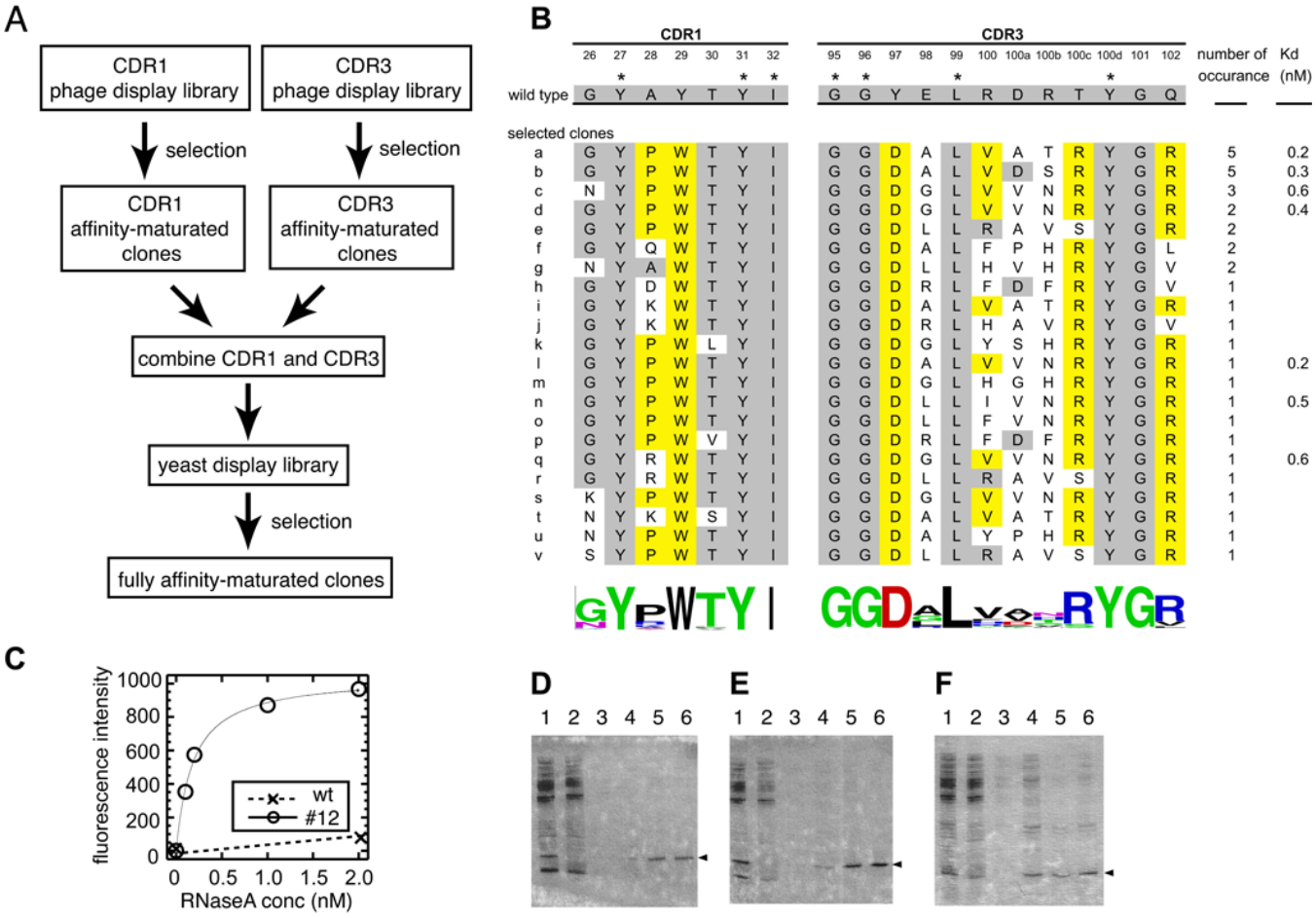

Figure 2. Affinity maturation of cAb-RN05 V_HH and characterization of a selected clone

(**A**) A flow chart showing the overall strategy of affinity maturation. (**B**) Amino acid sequences of the wild-type V_HH and selected mutants. K_d values for a subset of clones are also shown. Residues that were held constant in the combinatorial libraries are marked with an asterisk. Residues that are identical to the wild type are shaded in gray and those represent a new concensus are in yellow. Degrees of sequence diversity among the selected clones are shown using the WebLOGO respresentation.[41; 42] (**C**) Antigen titration curves for the wild type (crosses) and an affinity-matured clone *a* measured using yeast surface display. Average florescence intensity is plotted as a function of antigen concentration. (D-F) Specificity test using immobilized V_HH and E. coli lysate. Data for the wild-type V_HH (D), clone a (E) and human SUMO4 (F) are shown. Lane 1 shows the sample (a mixture of a protein with a His_6-tag and E. coli lysate) applied to the Ni-affinity column, lane 2 shows the flow-through fraction, lane 3 shows buffer wash fraction, and lanes 4-6 show elution fractions. Co-elution of the protein with a His_6-tag (marked with the triangles) and E. coli proteins, as seen in lane 4 of panel F indicate interactions of the protein with E. coli proteins.

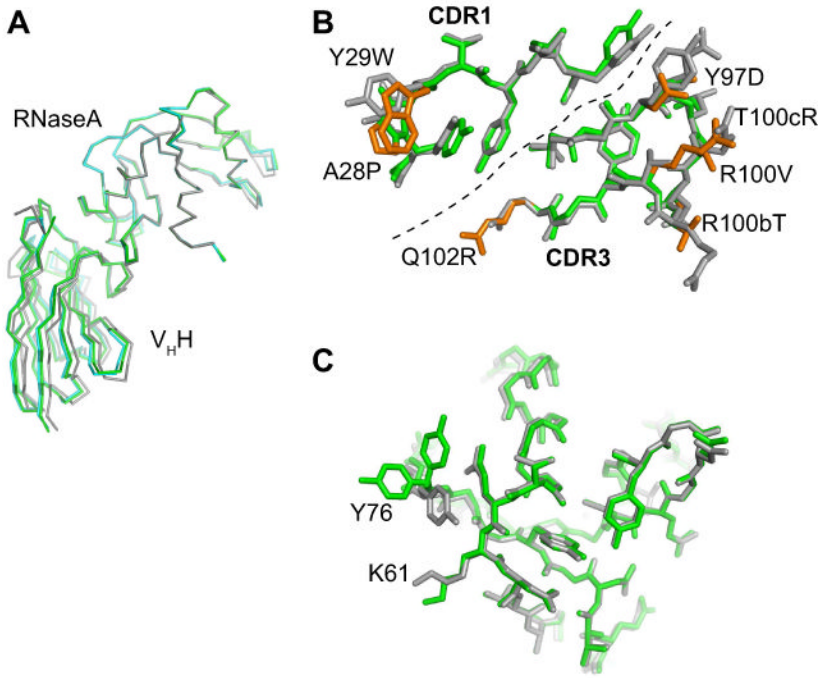

Figure 3. High-resolution x-ray crystal structures of wild-type and affinity-matured V_HHs in complex with RNaseA

(**A**) A comparison of the two structures after superposition of the RNaseA portion. The wild type is shown in gray and the mutant in green. (**B**) CDR1 and CDR3 residues of the wild-type (gray) and mutant (green) V_HHs. The dashed line divides CDR1 and CDR3. Mutations in the affinity-matured V_HH are shown in orange and labeled. (**C**) The epitope on RNaseA in the wild-type (gray) and mutant (green) complexes. Y76 has two conformers in the mutant structure.

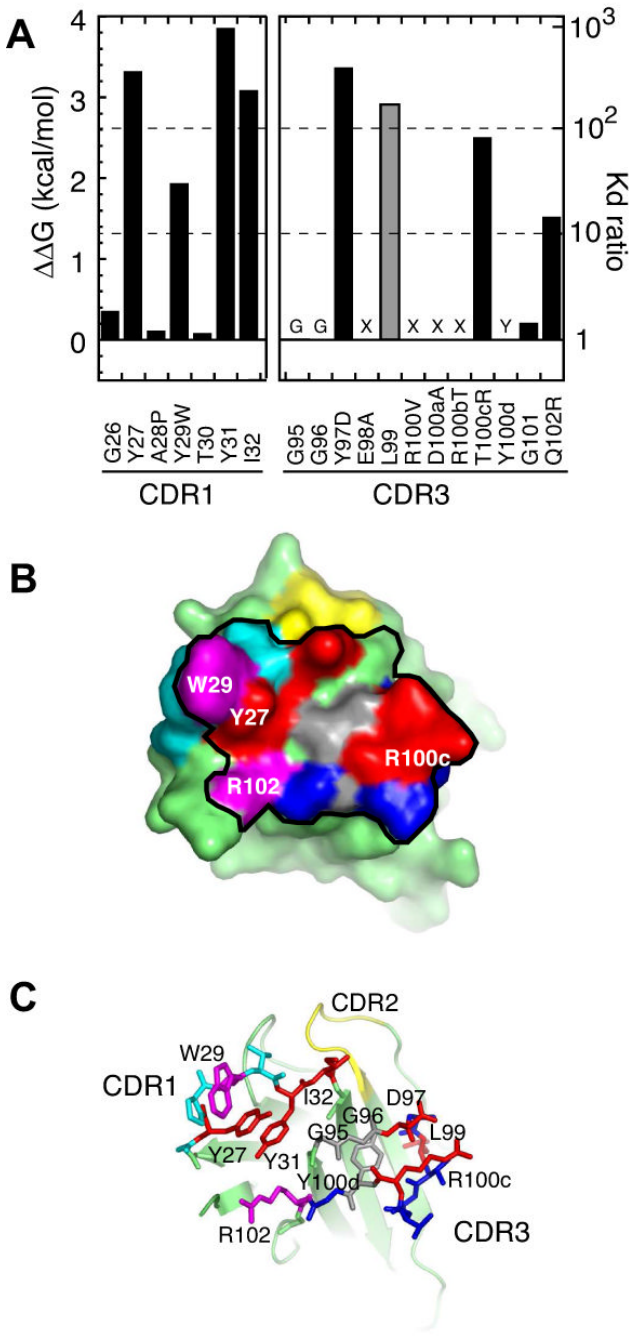

Figure 4. Energetic contributions of the paratope residues of the affinity-matured V_HH probed with Ala-scanning mutagenesis

(**A**) Effects of Ala substitution on antigen binding expressed in terms of ΔΔG (the left axis) and the ratio of K_d (= $K_d^{wild\text{-}type}/K_d^{mutant}$; the right axis). Positions 95, 96, 98, 100-100b and 100d were not tested because (i) Gly residue is critical (labeled with "G"), (ii) no sequence convergence was found in library screening (labeled with "X") or (iii) the Tyr is involved in scaffolding (labeled with "Y"). The value for L99, a noncontacting residue, is shown in gray. (**B**) The surface representation of the affinity-matured V_HH. The black line encloses residues within 5 Å of antigen atoms. Red and cyan surfaces are for residues where Ala substitution causes >100 fold and 10-100 fold increase in K_d, respectively. Gray surfaces indicate critical

residues that were not tested with Ala scanning. The remaining CDR1 and CDR3 residues are shown in light blue and dark blue, respectively. The four residues that expand the hot spot (W29, D97, R100c and R102) are labeled. (**C**) shows CDR1 and CDR3 residues as sticks and the V_HH scaffold as a cartoon model.

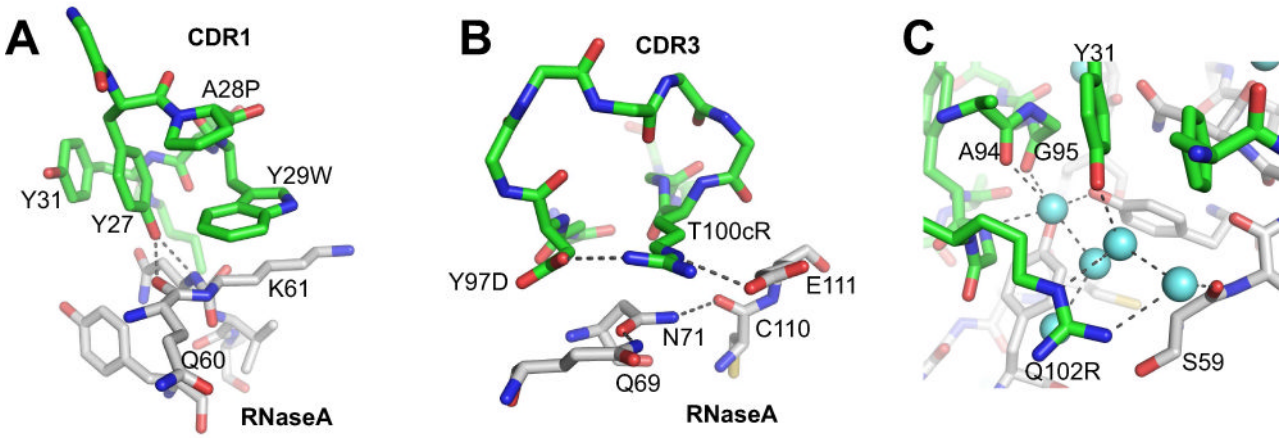

Figure 5. Molecular interactions involving mutated residues in the affinity-matured V_HH

The carbon atoms of V_HH are in green and those of RNaseA are in white. The nitrogen and oxygen atoms are in blue and red, respectively. (**A**) Interactions involving CDR1 residues. H-bonds are depicted with dashed lines. (**B**) Interactions involving triads of polar residues. (**C**) Interactions involving Q102R and interface water molecules (spheres in light blue). See the main text for more details.

Table 1

Interface characteristics of wild-type and affinity-matured cAb-RN05 and related proteins.

Ligand	K_d (nM)	ΔG (kcal/mol)	Epitope			Paratope			SC[2]	Planarity
			# of atoms[1]	Ligand Efficiency	Area (Å^2)	# of atoms[1]	Efficiency	Area (Å^2)		
cAb-RN05										
WT	23	10.3	46	0.22	589	47	0.22	613	0.76	2.13
matured	0.15	13.2	63	0.21	562	51	0.26	592	0.78	2.02
Other V_HH										
D3-L11 (1ZVY)	0.14	13.3	43	0.31	729	58	0.23	649	0.77	2.43
D2-L29 (1ZV5)	10	10.8	45	0.24	759	60	0.18	634	0.75	2.88
cAb-Lys3 (1JTT)	7.5	10.9	68	0.16	690	54	0.20	793	0.75	3.3
cAb-Lys2 (1RI8)	0.077	13.6	52	0.26	795	67	0.20	771	0.71	3.08
D2-L19 (1RJC)	3	11.5	59	0.19	741	60	0.19	644	0.79	2.88
D2-L24 (1ZVH)	80	9.5	63	0.15	528	51	0.19	593	0.82	1.65
cAb-Hu6 (1OP9)	0.7	12.3	45	0.27	556	43	0.29	577	0.76	1.61
Small molecule										
SP4206-IL2R (1PY2)	68.8	9.6	44	0.22	422	30	0.32	572	0.77	2.01

[1] Atoms within 4Å of the binding partner.

[2] Surface complementarity.[21]

J Mol Biol. Author manuscript; available in PMC 2008 November 2.

Table 2

X-Ray data collection and refinement statistics

	Wild-type complex	Affinity matured complex
Data collection[a]		
Space group	P 2_1	C 2 2 2_1
Cell dimensions		
a, *b*, *c* (Å)	40.13, 54.39, 48.04	70.03, 131.62, 137.16
α, *β*, *γ* (°)	90, 108.5, 90	90, 90, 90
Resolution (Å)[b]	50-1.38 (1.43-1.38)	50- 1.90 (1.97-1.90)
R_{merge}	0.05 (0.50)	0.10 (0.39)
$I/\sigma I$	56.3 (6.1)	17.4 (5.2)
Completeness (%)	99.6 (99.3)	99.1 (93.7)
Redundancy	14.1 (13.5)	7.1 (4.9)
Refinement		
Resolution (Å)	20-1.38	20 – 1.90
No. reflections	38,187	48,836
R_{work}/ R_{free}	15.7 / 18.9	18.5 / 21.7
No. residues		
V_HH	121	2×121
RNaseA	124	2×124
Ions	PO_4	2 SO_4
Water	366	357
B-factors		
V_HH	15.1	35.3
RNaseA	14.6	29.0
Ions	20.1	71.6
Water	27.6	37.2
R.m.s deviations		
Bond lengths (Å)	0.014	0.008
Bond angles (°)	1.74	1.15
Ramachandran plot statistics		
Most favored	90.3% (187)	91.5% (386)
Additional allowed	9.2% (19)	8.1% (34)
Generously allowed	0.5% (1)	0.5% (2)
Disallowed	0	0

[a] Two crystals were used to obtain the complete data set in case of the affinity matured complex. Only one crystal was used for wild-type complex.

[b] Data in parentheses correspond to two last resolution shells.