Published in final edited form as:

J Mol Biol. 2008 May 2; 378(3): 622–633.

Synthetic Antibody Libraries Focused Towards Peptide Ligands

Christian W. Coughlin¹, Juan C. Almagro², Mark Pogson³, Brent Iverson⁴, and George Georgiou⁵

¹ University of Texas at Austin, Institute for Cell and Molecular Biology, 2500 Speedway MBB 3.312, Austin TX, 78712, Ph: 512.791.9680, coughlin@utexas.edu

² Centocor R&D, Inc., 145 King of Prussia Rd., Radnor, PA 19087, Office: R3621, Ph (office): 610.651.6782, email: jalmagro@centocor.com

³ University of Texas at Austin, Institute for Cell and Molecular Biology, 2500 Speedway MBB 3.312, Austin TX, 78712, Ph: 512.232.4104

⁴ University of Texas at Austin, Institute for Cell and Molecular Biology and Department of Chemistry and Biochemistry, Campus Mail Code A5300, Austin TX, 78712, biverson@mail.utexas.edu

⁵ University of Texas at Austin, Institute for Cell and Molecular Biology and Department of Chemical Engineering, Austin TX, 78712, gg@che.utexas.edu

Abstract

Synthetic antibody libraries have proven immensely useful for the *de novo* isolation of antibodies without the need for animal immunization. Recently, focused libraries designed to recognize particular classes of ligands, such as haptens or proteins, have been employed to facilitate the selection of high affinity antibodies. Focused libraries are built using V regions encoding combinations of canonical structures that resemble the structural features of antibodies that bind the desired class of ligands and sequence diversity is introduced at residues typically involved in recognition. Here we describe the generation and experimental validation of two different scFv libraries that efficiently generate binders to peptides, a class of molecules that has proven to be a difficult target for antibody generation. First, a human anti-peptide library was constructed by diversifying a scaffold: the human V_H germline gene 3-23 which was fused to a variant of the human V_L germline gene A27, in which L1 was modified to encode the canonical structure found in anti-peptide antibodies. The sequence diversity was introduced into 3-23 (V_H) only, targeting for diversification residues commonly found in contact with protein and peptide antigens. Second, a murine library was generated using the antibody 26-10, which was initially isolated based on its affinity to the hapten digoxin, but also binds peptides and exhibits a canonical structure pattern typical of anti-peptide antibodies. Diversity was introduced in the V_H only using the profile of amino acids found at positions that frequently contact peptide antigens. Both libraries yielded binders to two model peptides, angiotensin and neuropeptide Y, following screening by solution phage panning. The mouse library yielded antibodies with affinities below 20 nM to both targets even though only the V_H had been subjected to diversification.

Correspondence to: George Georgiou.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Immunoglobulins; Igs; Hypervariable Loops; Framework; FR; Complementarity Determining Regions; CDR; Canonical Structures; phage display libraries; antibody engineering

Introduction

The discovery of antibodies for therapeutic, diagnostic, analytical or affinity chromatography purposes relies increasingly on the use of large combinatorial libraries encoding repertoires of antibody variants. Antibodies that recognize and bind to a target with the requisite affinity and specificity are isolated from a library by high throughput screening techniques such as phage, bacterial, yeast or ribosome display^{1; 2; 3}. Three distinct types of antibody libraries have been generated: libraries that sample the antibody diversity available in immunized animals⁴, libraries that mimic the diversity of the antibodies in non-immunized individuals³ and synthetic libraries². In the latter, the sequence of one or more antibody scaffolds is diversified so that it encodes all or a subset of the 20 natural amino acids at multiple positions known to be critical for antigen recognition. The theoretical diversity of synthetic antibody libraries in which multiple hypervariable loops are randomized far exceeds practical limitations in library construction and screening. Incorporating information on how the natural immune repertoire has evolved to recognize foreign antigens can facilitate the generation of more useful synthetic antibody libraries⁵.

The growing number of antibody structures^{6; 7; 8; 9} allows for a comprehensive correlation of antibody binding site topography with the various classes of recognized antigens. Binding site topography, in turn, is derived largely from hypervariable loop lengths and canonical conformations. Concave, grooved, or barrel-shaped binding sites are typically found in antibodies that bind to small molecules while flattened antigen-binding surfaces are involved in protein recognition. The binding site topographies of peptide binders lies somewhere between these two⁸. Persson, *et al.*¹⁰ recently reported the construction of a synthetic library biased towards the isolation of antibodies against small molecules by incorporating the structural and sequence features known to occur in many anti-hapten antibodies. However, to date, there are no reports describing the creation of libraries tailored to the isolation of anti-peptide antibodies.

Due to the conformational flexibility of peptides and the loss of entropy that occurs when they are bound by antibodies, the generation of high affinity antibodies to peptides can be challenging¹¹. Peptides are attractive therapeutic and diagnostic targets due to their varied physiological roles, from signaling molecules as hormones or neurotransmitters to secreted toxins^{12; 13}. Herein, we introduce a novel strategy to generate *de novo* synthetic antibody libraries focused for peptide-binding. By leveraging the canonical structural features of anti-peptide antibodies, it was hypothesized that libraries could be designed with a predisposition to bind peptides. We report the construction of two libraries based, respectively, on a human (Hu:anti-pep) and on a mouse (M:anti-pep) scaffold antibody. Both libraries have diversified V_H regions but invariant V_L regions. The scaffold of the Hu:anti-pep library is derived from well-characterized human germline V_H and V_L genes, with H3 and L1 loop length modifications to resemble the antigen-binding site topographies of antibodies binding peptides⁸. The diversity is limited at low-usage specificity determining residues (SDRs) and unlimited at high-usage SDRs in antibodies that interact with proteins and peptides¹⁴. The M:anti-pep library utilizes the well-expressed murine anti-digoxin 26-10 antibody scaffold that contains canonical structures that are similar to other anti-peptide antibodies⁸. The M:anti-pep library contains limited diversity at SDRs that recognize peptides¹⁴. Following screening by phage display, both libraries generated specific binders against two non-related peptides, with

antibodies isolated from M:anti-pep library having dissociation constants below 20 nM. The validated scFv libraries which efficiently generate peptide binders, as reported here, are likely to facilitate the isolation of high affinity antibodies to peptides of biochemical or therapeutic interest.

Results

Design and construction of the Hu:anti-pep library

Design and validation of the V_H component of the Hu:anti-pep library has been described elsewhere¹⁴. Similar to the M:anti-pep library described below, all diversified SDRs in the Hu:anti-pep library are located in the V_H domain, while the V_L domain is kept invariant (Figure 1). The V_H library was designed starting from the V_H human germline gene 3-23 recombined with the J_H4 minigene *via* an eight-residue H3 loop. Diversity was generated at a consensus of anti-protein and anti-peptide SDRs. High usage SDRs were varied with all 20 amino acids whereas SDRs that make contact less frequently were limited to Tyr, Asp, Ala, or Ser (by using oligonucleotides with the degenerate codon KMY¹⁴). Tyr and Asp play a dominant role in antigen recognition, while the Ala and Ser residues allow for space and conformational flexibility, functioning in an auxiliary role to facilitate contacts between Tyr or Asp and the antigen^{15; 16}. Position 93 of H3 was fixed to an Ala whereas position 94 was allowed to encode the more typical germline amino acid Arg¹⁷, in addition to the Lys found in 3-23. The H3 apical region (positions 95, 96-98) as well as positions 50 and 52 of H2 and position 31 of H1 encoded all 20 amino acids (NNK codon scheme) whereas positions 30, 32, 33, and 35 of H1, 52a-54, 56, and 58 of H2 encoded Tyr, Asp, Ala, or Ser.

The V_H library described above was cloned with a synthetic V_L . The latter was constructed by modifying the human germline V_K gene A27, a highly promiscuous light chain, capable of generating antigen-binding sites that accommodate a wide diversity of antigens¹⁸ and that is overexpressed *in vivo*^{19; 20; 21}. A27 codes for flat antigen-binding sites and displays a type 6 canonical structure in L1²². To modify it to encode a grooved antigen-binding site typical of anti-peptide antibodies, L1 from the germline B3 gene, which is a canonical structure type 3 (the longest canonical structure of L1), was grafted into A27 (Figure 2). The hybrid A27, called A27 modified (A27md), was fused to the J κ 1 region to create the complete V_L scaffold. The library was cloned into the pHEN vector¹⁹, resulting in 2×10^8 independent transformants. A total of 20 clones were picked at random and sequenced. All clones contained the insert and intended diversity with two clones each containing one stop codon.

Design and construction of the M:anti-pep library

For the murine library a different design was implemented (Figure 1). First, several commonly used, stable, and well-expressed antibodies were evaluated as potential scaffolds for the construction of a library biased towards peptide binders. Features of anti-peptide antibodies such as canonical structure patterns, loop lengths, SDRs and sequence composition were examined in candidate scaffold genes. Vargas-Madrado, *et al.* noted that 52% of the antibodies with the canonical structure pattern of 1-2-4-1-1 (H1-H2-L1-L2-L3) are peptide binders⁸. The canonical structure designations rely heavily on the HV loop lengths. In addition, we found that the dominant HV loop lengths of 46 anti-peptide antibodies in the database are as follows⁷: L1: 11 (35%) or 16 (55%); L2: 7 (100%); L3: 8 (9.5%) or 9 (88%); H1: 5 (85%); H2: 17 (72%); H3: 7 (22%), 8 (13%), or 12 (13%).

The mouse anti-hapten antibody, 26-10, which binds to the cardiac glucoside digoxin with a nanomolar dissociation constant²³, displays the 1-2-4-1-1 canonical structure pattern typically found in anti-peptide antibodies. Also, the loop lengths of 26-10 conform to those of anti-peptide antibodies: L1: 16; L2: 7; L3: 9; H1: 5; H2: 17; H3: 10 (Table S1). The 26-10 scFv is

well expressed in bacteria, it is highly soluble, stable, and displays a minimal propensity to form dimers or higher molecular weight aggregates in solution, making it well-suited for library selection experiments. For these reasons, even though the 26-10 antibody was elicited through immunization with digoxin, it was selected as the scaffold for construction of an anti-peptide library.

Analysis of the contact residues in eight antibodies that bind to small polypeptides⁹ and 18 anti-peptide antibodies⁵ was employed to identify SDRs that occur preferentially in anti-peptide antibodies. All 11 SDRs in the co-crystal of 26-10 with digoxin²⁴ correspond to SDRs found in anti-peptide antibodies in the aforementioned studies, making these sites candidates for sequence diversification. Positions with solvent-exposed side chains in 26-10 corresponding to SDRs in anti-peptide antibodies were also diversified. To simplify the construction of the library, diversity was introduced only in the V_H. No positions in V_L were diversified (Figure 1).

The positional frequencies of amino acids at all of the SDRs located in H1 and H2 of anti-peptide antibodies were determined by aligning 103 heavy chains available in the Kabat database²⁵. Amino acid frequencies for each SDR position were determined (Table 1). Subsequently, degenerate codons were designed for each SDR of H1 and H2 with an attempt to encode for the natural diversity of amino acids found at each SDR position. In some cases, two or more degenerate codons, and therefore two or more oligonucleotides, were necessary in order to encode most of the expected amino acid diversity (Tables 1 and 2) while eliminating termination codons and cysteines.

The length of the H3 loop was allowed to vary (n=8, 10, and 12) to reflect the length diversity of other anti-peptide antibodies. In H3, a consensus pattern for the four amino acid positions just prior to position 103 of Ala-Met-Asp-Tyr is often observed in natural antibodies²⁶. Similarly, Ala and Arg/Gly are conserved at positions 93 and 94, respectively^{27; 28}. These consensus sequence motifs are maintained in the H3 library design. Between four and nine positions in H3 were diversified in 26-10, depending on the length variation introduced. Primers of appropriate lengths were designed using the NNS coding scheme (S= G or C) which encodes for all 20 amino acids and also one stop codon. Allowing all 20 amino acids in the randomized positions of this H3 loop of the library could enable a better sample of sequence space and result in higher affinity clones²⁶.

Following gene construction and ligation into pAK200 vector⁴, a total of 5.1×10^8 independent transformants were obtained. 12 clones, picked at random, were sequenced and found to contain inserts with the expected diversity. The length diversity of loop H3 was also well represented (n=8, 5 out of 12; n=10, 3 out of 12; n=12, 4 out of 12).

Selections of scFvs with 2 different peptide targets

The Hu:anti-pep and M:anti-pep libraries were screened by phage display against the peptides human angiotensin (AT, 10 amino acids, M.W. 1296 Da) and human neuropeptide Y (NPY, 36 amino acids, M.W. 4271 Da). Both peptides were synthesized with a biotin moiety at the N-terminus. Three rounds of panning were performed in each case, except for the M:anti-pep library, which was panned against NPY for five rounds. Enrichment was monitored by polyclonal phage enzyme-linked immunosorbent assay (ELISA) (Figure S1).

Individual clones from the last round of selection were randomly chosen for additional screening by monoclonal phage ELISA. Specifically, phage antibodies from 96 individual clones selected at random from the last round were assayed for binding to either the selection peptide or an unrelated peptide as a control (Supplementary Figure 2). Following three rounds of selection, 71% of the screened clones from the Hu:anti-pep library panned against NPY

showed specificity to NPY, while 87% of the clones from the Hu:anti-pep panned against AT showed specificity to AT. For the M:anti-pep library, 79% of the clones panned against AT showed specific binding after two rounds and 96% after three rounds. A total of 99% of the clones from the fifth round of selection of M:anti-pep library against NPY showed specific binding. Clones exhibiting a binding signal to noise ratio greater than 10 by monoclonal phage ELISAs were sequenced (Tables 3 and 4). 15 clones were sequenced from the both the AT and NPY panning of the Hu:anti-pep library revealing two distinct clones for each peptide. For the M:anti-pep library, after two rounds of screening against AT, nine of 12 clones were distinct. After the third round, sequencing of twelve clones yielded only two distinct sequences, one of which, M:ATB5, was already present in round two. Interestingly, 20 of the 24 total M:anti-pep clones specific for AT contained the same H2 DNA sequence as the parental 26-10 scFv. In contrast, the round 5 phage population of M:anti-pep library screened with NPY exhibited a high degree of diversity and contained 13/15 distinct clones (Table 4).

Based on the monoclonal ELISA sequencing results, human and mouse clones specific for either AT or NPY were chosen for further characterization. Proteins were expressed in soluble scFv form in *E. coli* and purified by immobilized metal affinity chromatography (IMAC) followed by gel filtration to remove dimers or higher molecular weight aggregates. All nine scFv proteins expressed well with seven of nine resulting in yields of purified, monomeric protein >1 mg/L (Table 5). The kinetics of antigen binding were determined by surface plasmon resonance on a BIACore 3000 instrument. Briefly, the biotinylated peptides were immobilized on a streptavidin chip at a surface density sufficient to give 30 (AT) or 75 (NPY) RUs, respectively, and the kinetics of association and dissociation were measured. The antibodies isolated from the Hu:anti-pep library, in general, exhibited lower affinities relative to the antibodies from the M:anti-pep (Table 5). Specifically, three out of four antibodies from the Hu:anti-pep library exhibited dissociation constants (K_D values) for their ligands in the micromolar range. In contrast, with the exception of the anti-AT antibody fragment M:ATE5, five of six antibodies from the M:anti-pep library displayed K_D values in the 10-100 nM range. This is particularly noteworthy considering that the V_L domain of the M:anti-pep library had not been subjected to diversification.

Discussion

Examination of the available structures of anti-peptide antibodies has revealed that they display similar topographical features, based mostly on loop lengths and canonical structures. It is reasonable to propose that the similar topography facilitates recognition of a general class of antigens, in this case peptides^{8; 9}. Shared sequence-based features include SDRs common to peptide binders and also the amino acid preferences displayed at these SDR positions.

The available information on the structure and sequence features of anti-peptide antibodies was exploited to construct human (Hu:anti-pep) and mouse (M:anti-pep) anti-peptide libraries. The V_H component of the Hu:anti-pep repertoire has been validated elsewhere²⁰. The V_K germline chain used to synthesize the A27md, is capable of recognizing a wide variety of antigens^{14; 18} and has also been validated in several other phage display pannings^{10; 15; 16}. To resemble the overall structure of antibodies that recognize peptides, an insertion of five residues was introduced at the tip of L1 in A27md. This insertion, combined with a short H3 loop, generates a grooved antigen-binding site (Figure 1) suitable for mid-sized antigens like peptides.

In contrast to the Hu:anti-pep library, the 26-10 murine antibody used as a scaffold for the M:anti-pep library, was isolated from a hybridoma that had undergone somatic hypermutation²⁹. The anti-digoxin 26-10 scaffold was selected primarily based on the structural features it shares with other anti-peptide antibodies. These shared structural features

include: (i) a peptide-specific canonical structure pattern (1-2-4-1-1) and (ii) loop lengths nearly identical to those of anti-peptide antibodies. Additional similarities between 26-10 and anti-peptide antibodies were observed in the SDR usage, indicating that the side chains at the putative SDRs of 26-10 are well positioned to interact with peptide antigens. 26-10 is also known to bind to peptides³⁰, perhaps because of its anti-peptide like binding pocket topography and SDR side chain orientation. Other desirable features of the 26-10 antibody as a scaffold for synthetic library construction include its high expression level in *E. coli* and its thermodynamic stability³¹.

Both libraries incorporate amino acid diversity in the V_H chain only. The canonical structure patterns of the V_H chains for both libraries are identical. The differences between the canonical structures in the V_L chains are subtle. L1 in the Hu:anti-pep library is one amino acid longer to yield the 3-1-1 canonical structure as compared to the 4-1-1 arrangement in M:anti-pep. While the 17 amino acid long L1 loop of the Hu:anti-pep scaffold is not typically associated with peptide binders, an examination of the canonical structures available in the PDB and a model of the library (Figure 1) indicates that the 3-1-1 binding pocket is likely concave enough to accommodate peptides. Also, the relatively long length of CDRL1 in both libraries could prevent interactions between flat continuous epitopes characteristic of protein antigens and the residues at the bottom of the antigen-binding site, thus focusing both libraries for the recognition of peptides.

The M:anti-pep library incorporated only the naturally observed diversity by aligning known anti-peptide antibody sequences. Residues were targeted for diversity based on two considerations: their use as SDRs in anti-peptide antibodies and in 26-10, and their surface exposure. Since the targeted diversity was based on that of existing antibodies, the number of structurally deficient clones should be lower³². Although not buried, position H47 is occupied by Trp in 99% of reported anti-peptide antibodies, whereas it is a Tyr in 26-10. For this reason, the diversity at that position in the M:anti-pep library included Tyr in addition to Trp. Length diversity was included in the M:anti-pep library for H3 to reflect the range of H3 lengths of natural anti-peptide antibodies.

The number of randomized positions in the Hu:anti-pep library (17) was lower than in the M:anti-pep library (21-25, depending on H3 length). Additionally, the Hu:anti-pep had a simpler diversification scheme: all 20 amino acids were incorporated at “high use” SDRs and only 4 amino acids at “low use” SDRs. Varying the positions that correspond to frequently used SDRs to all 20 amino acids enabled a large number of possible amino acids to be explored only at the positions critical to peptide recognition. The SDR positions identified as infrequently used SDRs in other anti-peptide antibodies^{5; 9} had diversity restricted to Tyr, Asp, Ala, or Ser in order to limit the total sequence space of the library. This limited diversity at the low usage SDRs has been validated for high affinity interactions with protein antigens^{14; 16}. Due to the limited role the side chains at these positions play in antigen recognition, restricting the diversity to these four amino acids was anticipated to have only minimal adverse effects on the activity of library clones.

Anti-peptide antibodies to AT and NPY were isolated from each of the two libraries following phage panning in solution. AT and NPY are representative of peptides having different lengths and solution structures. Although AT is short, at only 10 amino acids, a solution structure³³ reveals a short β -sheet in the middle of the peptide. At 21.8 Å in length, AT is only slightly longer than digoxin, the hapten recognized by the M:anti-pep scaffold antibody 26-10 (~19 Å). The solution structure of the larger peptide, NPY (36 amino acids, M.W. 4271 Da), with a length of 51 Å³⁴, shows a six turn α -helix at its C-terminus. Thus, AT is close to the size range of the 26-10 binding peptides (15mers around 1800 Da, 26mers around 2900 Da) isolated by Ball *et al.*³⁰, whereas NPY is much longer.

Isolated clones from each library bound specifically only to their target peptide. Two high affinity clones, M:ATB5 and M:ATG11 (K_D values of 14.2 nM and 52.6 nM, respectively) were isolated from the M:anti-pep library after two (M:ATB5) and three (M:ATG11) rounds of selection against AT. Interestingly, three AT-binding clones from the M:anti-pep library had the same H2 sequence, which is identical to the H2 sequence of the scaffold. Since this sequence is present in both the aforementioned high affinity clones and also in the low affinity clone M:ATE5, it is possible that it was isolated because it was slightly overrepresented in the library as a result of PCR contamination during the construction of the library from the wild-type H2 and may not reflect the optimal sequence for antigen-binding. On the other hand we note that the two NPY-specific clones did not contain the 26-10 H2 loop. Hence, it is possible that its selection in the M:ATB5, M:ATE5 and M:ATG11 antibodies reflects a possible structural role in stabilizing the binding pocket or enhancing specificity for smaller peptides such as AT.

The sequences of the two higher affinity M:anti pep-AT clones, M:ATB5 and M:ATG11 are identical except for the H3 region. The H3 length of M:ATG11 ($K_D=52.6$ nM) is 8 amino acids, while H3 of M:ATB5 ($K_D=14.2$ nM) is 10 amino acids. The side chains at the apex of H3 are not accessible to the binding pocket²⁴ when the H3 loop length is 10 amino acids or longer, rendering any interaction between the apical positions of H3 of M:ATB5 and AT unlikely. A key SDR in the parental 26-10 antibody, H100Trp is replaced with smaller amino acids, namely Ala in M:ATB5 and Thr in M:ATG11 at position H98 which is equivalent to position H100 when H3 length is eight. The presence of smaller amino acids could enlarge the binding pocket, allowing the AT peptide to bind in an orientation analogous to that of digoxin in 26-10. The sequence of the low affinity M:anti-pep clone, M:ATE5, is different from the two higher affinity clones in H1 as well as in the H3 loop. The H3 loop length of M:ATE5 is eight amino acids and the sequence is almost identical to that of the higher affinity M:ATB5. The Lys at position H97 of M:ATE5 is likely oriented towards the top of the binding pocket and could hinder entry of AT into the binding pocket compared to the much smaller Val or Gly at this position in M:ATG11 or M:ATB5, respectively. Consistent with this notion, M:ATE5 exhibits a slower association rate.

Binders to NPY from the M:anti-pep library were isolated after five rounds of panning. Although the sequence diversity in round five was still high, additional rounds of screening resulted in a low polyclonal phage ELISA signal indicating de-enrichment (data not shown). Two clones showing the highest ELISA signal, M:NPYSol1 and M:NPYSol2, exhibited K_D values of 75.2 and 18.3 nM, respectively. The respective amino acid sequences are distinct, but both exhibit an eight amino acid H3 loop length. A shorter H3 loop may serve to form a more shallow grove capable of accommodating the helical structure of NPY⁸. The solvent exposed apex of the H3 loop in M:NPYSol2 has two Trp residues. Such solvent exposed hydrophobic patches are often destabilizing in antibodies³⁵. However, because of the short H3 of M:NPYSol2, these two Trp residues might actually be part of the binding pocket giving the hydrophobic side chains an opportunity to interact with NPY.

After three rounds of panning of the Hu:anti-pep library we obtained two distinct clones for each peptide target. The two clones specific for AT, Hu:AT1 and Hu:AT2 exhibited only micromolar K_D values. Interestingly, one of the clones showed a shorter H3 loop than expected based on the design. Despite differences in the length of H3 among the two clones, the k_{on} s were very similar: Hu:AT1 $k_{\text{on}}=0.0139$ s⁻¹ and Hu:AT2 $k_{\text{on}}=0.0112$ s⁻¹. The two NPY binders isolated from the Hu:anti-pep library differed markedly in their affinities. For clone Hu:NPY1, an unusually slow association rate ($k_{\text{on}}=5.09 \times 10^3$) was primarily responsible for the low overall affinity ($K_D=4.1$ μ M) relative to the second clone, Hu:NPY2, which exhibited a $K_D=48.6$ nM. The slow association rate constant indicates that Hu:NPY1 may have to undergo a rate limiting conformational change in order to bind to the antigen. Alternatively, the two aspartic acids at

positions 31 (in the H1 loop) and 58 (H2) could be negatively impacting antigen binding. Recent computational studies have underscored the significance of electrostatic interactions in antibody:antigen recognition³⁶. In addition Hu:NPY2 contains a Trp at position 31 (in H1) compared to a Ser in the lower affinity Hu:NPY1 antibody. The sequence differences of the H3 loops of these two antibodies are relatively minor. The most notable difference is the presence of a His at position 98 of Hu:NPY1 compared to Gly in Hu:NPY2. We note that Hu:NPY1 may have been selected in part because it is very well expressed resulting in a yield of 4.8 mg/L of purified protein, more than 15-fold better than what was obtained with Hu:NPY2 (0.3 mg/L).

Possibly due to the flexible nature of peptides in the unbound state and the loss of entropy in the antibody-bound state¹¹, previous attempts to isolate high-affinity anti-peptide antibodies from synthetic libraries has led to mixed results. The presence of relatively few hits against peptide ligands compared to protein antigens suggests a lack of peptide specificity in screening highly complex synthetic libraries³⁷, resulting in isolated anti-peptide antibodies with K_D values $> 1\mu\text{M}$ ^{37; 38}. Anti-peptide antibody affinities ranging from 67-420 nM were reported for a library constructed by incorporating natural CDR diversity within a single scaffold³⁹. On the other hand, as might be expected, immune repertoire libraries have resulted in antibodies with far greater affinity (K_D values in the picomolar range) to peptides⁴⁰. Most notably, affinity maturation using ribosome display of an antibody isolated from an immune repertoire library has even resulted in an anti-peptide antibody with a K_D of 1pM⁴¹.

Due to the *de novo* nature of the M:anti-pep and Hu:anti-pep libraries, there were no corresponding general purpose libraries against which to measure the peptide-binding efficiency of the libraries described here. It is important to keep in mind that in both libraries the light chain was not subjected to diversification and consequently most, if not all, of the binding energy defining the specificity for AT or NPY likely arose from contacts with SDRs in the V_H . Despite the invariant natural or rationally designed V_L chain of these libraries, medium to high affinity, specific antibodies were generated from each library. In the case of M:anti-pep library nanomolar affinity antibodies were isolated against the two non-related peptides. Thus, encouraged by our early success in creating a fertile anti-peptide synthetic antibody library, work is currently underway to design next generation libraries that include V_L diversification.

Materials and Methods

Library Construction

Hu:anti-pep Library—The scaffold of the Hu:anti-pep library comprises the V_H of 3-23 (DP-47)/J_H4¹⁹, and the V_L of A27md and J_κ1. The construction of the V_H repertoire was described earlier¹⁴. A gene encoding the V_H was purified, digested with *Xho*I and *Sfi*I, and ligated into the phagemid vector pHEN1- V_L A27md/J_κ3 as previously described¹⁴. The library DNA was amplified as above and following electroporation transformants were selected by plating on plates containing 100 $\mu\text{g/L}$ ampicillin instead of chloramphenicol. 2×10^8 independent transformants were obtained.

M:anti-pep Library—The anti-digoxin scFv 26-10^{23; 29} was first cloned into the pAK200 vector⁴. TAA stop codons were introduced into each of the three V_H HV loops and the resulting gene, “scFv 26-10 stop” used as the template for the construction of the library. The stop codons are repaired during the library construction. Six or 10 putative SDR positions in H1 and H2, respectively, were diversified, and between 4-9 positions in H3 were diversified depending on the length variation (n) introduced (n = 8, 10, or 12).

Diversity was introduced by separate PCR amplification of each of the three CDRs in the V_H using oligonucleotide primers containing degenerate codons using the codon diversification scheme proposed by Sidhu, *et al*³² (Tables 1 and 2). Length and sequence diversity of H3 was generated by three separate PCR reactions using oligonucleotides encoding eight, ten, or twelve amino acids. H1 was amplified using an upstream vector specific primer and two reverse library primers, H2 was amplified using six forward library primers and 1 reverse primer, and each of the H3 loops was amplified using two forward library primers and one downstream vector specific reverse primer that also amplified the V_L region. High Fidelity polymerase (Roche Diagnostics, Indianapolis, IN, USA) was used in all PCR reactions to limit mutations to those encoded by the oligonucleotides. Each CDR fragment was agarose-gel purified and extracted using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). The respective fragments were assembled into the complete scFv by overlap-extension PCR. This reaction included 150ng of DNA from each of the genes having diversified CDR regions region (*i.e.* H1, H2, H3₈, H3₁₀ and H3₁₂). The full length scFv fragment was PCR purified using the QIAquick PCR Purification Kit and digested with *Sfi*I (New England Biolabs, Ipswich, MA, USA) at 50°C for six hours, gel purified and ligated, using T4 ligase (New England Biolabs) overnight at 16°C, into the complementary sites of *Sfi*I-cut phagemid, pAK200⁴ using a molar ratio of 2:1 (insert to vector). Ligation reactions were desalted on nitrocellulose (Millipore, Billerica, MA, USA) and 200ng of DNA were electroporated into *E. coli* XL1-Blue (Stratagene, La Jolla, CA, USA) in 10 separate reactions, recovered, plated on TYE-2% glucose supplemented with 34µg/mL chloramphenicol and incubated overnight at 30°C⁴; ¹⁹ resulting in library of 5.1×10⁸ independent transformants.

Degenerate DNA is represented using the IUB code: N=A,C,G, or T; V=G,A, or C; B=G,T, or C; H=A,T, or C; D=G,A, or T; K=G or T; S=G or C; W=A or T; M=A or C; Y=C or T; R=A or G.

Phage Panning

Library cells were scraped and used to inoculate 50mL 2×YT-2% glucose containing chloramphenicol resulting in an OD₆₀₀ of 0.1. The cells were grown at 30°C to OD₆₀₀=0.6 and infected with a 100-fold concentration of VCS-M13 helper phage (Stratagene, La Jolla, CA, USA). Subsequently, the cells were incubated at 37°C for one hour, spun down, and resuspended in 100mL of 2×YT containing 0.1% glucose, chloramphenicol or ampicillin as above and kanamycin at 50 µg/L and incubated overnight at 30°C. Phage particles were harvested from the culture supernatant following two rounds of precipitation with 20% PEG₆₀₀₀/2.5M NaCl (Tomlinson Protocol, Geneservice, UK), stored at 4°C and panned within 18 hours.

Panning was performed in solution with biotinylated-human angiotensin I (AT; sequence: DRVYIHPFHL) and biotinylated human neuropeptide Y (NPY; sequence: YPSKPDNPGEDAPAEDMARYYSALRHYINLITRQRY) (Genscript, Piscataway, NJ, USA) using streptavidin-coated paramagnetic beads (Invitrogen, Carlsbad, CA, USA) essentially as described⁴². Prior to panning the phage with 10µM biotin-peptide, non-specific phage antibodies were depleted on 100µL naked streptavidin beads without peptide. The remaining phage antibodies were then incubated with 10µM biotin-peptide in 1 mL PBS before the phage antibody-biotin-peptide complexes were captured with 100µL streptavidin beads. The beads were washed vigorously and bound phage were eluted by incubating the beads with 1mL 100mM triethylamine (TEA) for 10 min while rotating. The beads were separated and the TEA-phage mixture was added immediately to 0.5 mL 1M Tris, pH 7.4 for neutralization. 9mL of *E. coli* XL1-Blue (OD₆₀₀= 0.7) culture were infected, the cells were plated on selective media as above, colonies were scraped and used to inoculate liquid cultures for phage preparation as above. The rescued phage particles were used in subsequent selection rounds.

Phage ELISAs

To evaluate the enrichment of selected phage pools the binding of polyclonal phage populations to the target peptide was monitored by polyclonal ELISA. Costar 3590 (Corning) plates were coated overnight⁴² with 100 μ L/well biotin-BSA (2 μ g/mL in PBS) (Sigma-Aldrich, St. Louis, MO, USA) at 4°C and washed three times with PBS-0.1% Tween20 (EMD) between reagents and samples. 100 μ L/well streptavidin (a 10 μ g/mL solution in PBS-0.5% gelatin, Jackson Laboratories) was then added for one hour at room temperature followed by 100 μ L/well of biotinylated-peptide (2 μ g/mL in PBS) for one hour at room temperature. The plates were blocked with 200 μ L/well PBS-2% milk for one hour. 100 μ L/well of normalized titers of phage antibodies from each round of panning were added to the plate in PBS-2% milk, in triplicate. After one hour at room temperature bound phage was detected with 100 μ L/well of anti-M13-horseradish peroxidase (HRP) conjugate (1:5000) (Amersham-Pharmacia Biosciences, Piscataway, NJ, USA) followed by development with 50 μ L/well TMB substrate (Dako) and quenching with 50 μ L/well 1N sulfuric acid. Plates were read at OD₄₅₀ and the signals from background binding were subtracted from the peptide binding.

To analyze individual clones, randomly selected clones from rounds of panning showing maximal enrichment by polyclonal ELISA were cultured overnight in microtiter plates (Tomlinson Protocol, Geneservice, UK). 50 μ L of phage supernatant from each well was collected and added to ELISA plates containing either the target peptide or an unrelated peptide. The plates were coated and developed as described above for the polyclonal analysis. A ratio of the absorbance reading for each clone's ability to bind to the target peptide and to the unrelated peptide was determined.

Expression of soluble scFvs

Infected cells collected from the last round of panning of the M:anti-peptide library as described, were cultured in 100mL TB (Difco) containing 2% glucose and chloramphenicol to an OD₆₀₀ of 1.5. Plasmid DNA was prepared, digested with *Sfi*I, gel purified, ligated into pAK400⁴ and the DNA was transformed into *E. coli* Jude-1⁴³ and plated on TYE with 2% glucose and chloramphenicol. Cloning into pAK400 allows expression of scFvs with a C-terminal His tag without fusion to the pIII.

Phage from the isolated clones of the Hu:anti-pep library were infected into HB2151 *E. coli*, an F⁺, non-suppressor strain (Maxim Biotech, Rockville, MD, USA) to promote recognition of the amber codon between the scFv coding region and gene III, enabling expression of His-tagged scFv from pHEN without the pIII fusion.

Selected clones from both libraries were grown in 400 mL TB containing 2% glucose, the appropriate antibiotic and one drop of anti-foam C (Sigma-Aldrich) at 25°C. After 3 hours, 1mM IPTG was added and incubation continued for an additional 4 hours. The cells were pelleted (5000g $\times$ 10min), and the periplasmic fraction was isolated. Briefly, the cell pellet was resuspended in 10mL ice-cold 100mM Tris-HCl, pH 7.4 containing 0.75M Sucrose and 1mL 10mg/mL lysozyme (Sigma-Aldrich). 20mL of 1mM EDTA was added, the cells were incubated for 15 minutes and then 1.4mL 0.5M MgCl₂ was added for another 15 min. The periplasmic fraction was separated from the spheroplasts by centrifuging the lysed mixture at 12,000g for 30 min. An equal volume of immobilized metal affinity chromatography buffer (IMAC; 20mM Na₂HPO₄, 500mM NaCl, 10mM imidazole, 0.1% Tween-20, pH 7.4) was added followed by 0.5mL washed Ni-NTA Agarose (Qiagen) for 1 hour. The resin was washed twice in IMAC buffer and elution was carried out with 4mL 250mM imidazole in IMAC buffer. Monomeric scFv was further purified by size exclusion gel filtration chromatography on a Superdex 200 column (Amersham-Pharmacia Biosciences, Uppsala, Sweden).

BIACore Analysis

Antigen binding kinetics of monomeric scFvs were determined *via* surface plasmon resonance analysis using a BIAcore 3000 (BIAcore-GE Healthcare, Piscataway, NJ, USA). Biotinylated peptides were captured on SA sensor chip with the reference cell containing streptavidin without peptide. Five different concentrations of each scFv were injected in triplicate over the sensor chips at a flow rate of 30 μ L/min for 2 min for k_a measurements. k_d was assessed during a 10 min dissociation stage. The chip was regenerated by a one min injection of 10mM glycine, pH 2.0 followed by a one minute injection of HBS-EP at a flow rate of 50 μ L/min. Response curves were generated by subtracting the response of the reference cell from that of the experimental cell and the data were analyzed using BIAevaluation v. 3.2. Binding kinetics were calculated by simultaneously fitting the k_a and k_d data to a 1:1 Langmuir binding model. Only data fits with a χ^2 value less than 2.5 were used.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported partially by a grant from ONR.

References

1. Mazor Y, Van Blarcom T, Mabry R, Iverson BL, Georgiou G. Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli*. *Nat Biotechnol* 2007;25:563–5. [PubMed: 17435747]
2. Sidhu SS, Fellouse FA. Synthetic therapeutic antibodies. *Nat Chem Biol* 2006;2:682–8. [PubMed: 17108986]
3. Hoogenboom HR. Selecting and screening recombinant antibody libraries. *Nat Biotechnol* 2005;23:1105–16. [PubMed: 16151404]
4. Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, Plückthun A. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J Immunol Methods* 1997;201:35–55. [PubMed: 9032408]
5. Almagro JC. Identification of differences in the specificity-determining residues of antibodies that recognize antigens of different size: implications for the rational design of antibody repertoires. *J Mol Recognit* 2004;17:132–43. [PubMed: 15027033]
6. Wilson IA, Stanfield RL. Antibody-antigen interactions: new structures and new conformational changes. *Curr Opin Struct Biol* 1994;4:857–67. [PubMed: 7536111]
7. Collis A, Brouwer B, Martin A. Analysis of the antigen combining site: correlations between length and sequence composition of the hypervariable loops and the nature of the antigen. *J Mol Biol* 2003;325:337–354. [PubMed: 12488099]
8. Vargas-Madrado E, Lara-Ochoa F, Almagro JC. Canonical structure repertoire of the antigen-binding site of immunoglobulins suggests strong geometrical restrictions associated to the mechanism of immune recognition. *J Mol Biol* 1995;254:497–504. [PubMed: 7490765]
9. MacCallum RM, Martin ACR, Thornton JM. Antibody-antigen Interactions: Contact Analysis and Binding Site Topography. *J Mol Biol* 1996;262:732–745. [PubMed: 8876650]
10. Persson H, Lantto J, Ohlin M. A focused antibody library for improved hapten recognition. *J Mol Biol* 2006;357:607–20. [PubMed: 16445941]
11. Zahnd C, Spinelli S, Luginbühl B, Amstutz P, Cambillau C, Plückthun A. Directed in Vitro Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. *J Biol Chem* 2004;279:18870–18877. [PubMed: 14754898]
12. Stresser DM, Kupfer D. Monospecific antipeptide antibody to cytochrome P-450 2B6. *Drug Metab Dispos* 1999;27:517–25. [PubMed: 10101147]

13. Sobel JH, Trakht I, Pileggi N, Qi Wu H. Antipeptide monoclonal antibodies to defined fibrinogen Aalpha chain regions: anti-Aalpha 487-498, a structural probe for fibrinogenolysis. *Blood* 1998;91:1590–8. [PubMed: 9473224]
14. Almagro JC, Quintero-Hernandez V, Ortiz-Leon M, Velandia A, Smith SL, Becerril B. Design and validation of a synthetic VH repertoire with tailored diversity for protein recognition. *J Mol Recognit* 2006;19:413–22. [PubMed: 16881059]
15. Fellouse FA, Barthelemy PA, Kelley RF, Sidhu SS. Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. *J Mol Biol* 2006;357:100–14. [PubMed: 16413576]
16. Fellouse FA, Wiesmann C, Sidhu SS. Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. *Proc Natl Acad Sci U S A* 2004;101:12467–72. [PubMed: 15306681]
17. Matsuda F, Ishii K, Bourvagnet P, Kuma K, Hayashida H, Miyata T, Honjo T. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J Exp Med* 1998;188:2151–62. [PubMed: 9841928]
18. Rojas G, Almagro JC, Acevedo B, Gavilondo JV. Phage antibody fragments library combining a single human light chain variable region with immune mouse heavy chain variable regions. *J Biotechnol* 2002;94:287–98. [PubMed: 11861087]
19. Griffiths A, Williams S, Hartley O, Tomlinson I, Waterhouse P, Crosby W, Kontermann R, Jones P, Low N, Allison T, Prospero T, Hoogenboom H, Nissim A, Cox J, Harrison J, Zaccolo M, Gherardi E, Winter G. Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J* 1994;13:3245–3260. [PubMed: 8045255]
20. Ohlin M, Borrebaeck CA. Characteristics of human antibody repertoires following active immune responses in vivo. *Mol Immunol* 1996;33:583–92. [PubMed: 8760269]
21. de Wildt RM, Hoet RM, van Venrooij WJ, Tomlinson IM, Winter G. Analysis of heavy and light chain pairings indicates that receptor editing shapes the human antibody repertoire. *J Mol Biol* 1999;285:895–901. [PubMed: 9887257]
22. Tomlinson IM, Cox JP, Gherardi E, Lesk AM, Chothia C. The structural repertoire of the human V kappa domain. *Embo J* 1995;14:4628–38. [PubMed: 7556106]
23. Huston JS, Levinson D, Mudgett-Hunter M, Tai MS, Novotny J, Margolies MN, Ridge RJ, Brucoleri RE, Haber E, Crea R, et al. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc Natl Acad Sci U S A* 1988;85:5879–83. [PubMed: 3045807]
24. Jeffrey PD, Strong RK, Sieker LC, Chang CY, Campbell RL, Petsko GA, Haber E, Margolies MN, Sheriff S. 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc Natl Acad Sci U S A* 1993;90:10310–4. [PubMed: 8234291]
25. Kabat EA, Wu TT. Attempts to locate complementarity-determining residues in the variable positions of light and heavy chains. *Ann N Y Acad Sci* 1971;190:382–93. [PubMed: 5290024]
26. Lee CV, Liang WC, Dennis MS, Eigenbrot C, Sidhu SS, Fuh G. High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. *J Mol Biol* 2004;340:1073–93. [PubMed: 15236968]
27. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Conformations of the third hypervariable region in the VH domain of immunoglobulins. *J Mol Biol* 1998;275:269–94. [PubMed: 9466909]
28. Shirai H, Kidera A, Nakamura H. H3-rules: identification of CDRH3 structures in antibodies. *FEBS Lett* 1999;455:188–97. [PubMed: 10428499]
29. Hunter MM, Margolies MN, Ju A, Haber E. High-affinity monoclonal antibodies to the cardiac glycoside, digoxin. *J Immunol* 1982;129:1165–72. [PubMed: 6179995]
30. Ball WJ Jr, Wang Z, Malik B, Kasturi R, Dey P, Short MK, Margolies MN. Selection of peptidic mimics of digoxin from phage-displayed peptide libraries by anti-digoxin antibodies. *J Mol Biol* 2000;301:101–15. [PubMed: 10926495]
31. Ribnicky B, Van Blarcom T, Georgiou G. A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. *J Mol Biol* 2007;369:631–9. [PubMed: 17462668]

32. Sidhu SS, Li B, Chen Y, Fellouse FA, Eigenbrot C, Fuh G. Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. *J Mol Biol* 2004;338:299–310. [PubMed: 15066433]
33. Spyroulias GA, Nikolakopoulou P, Tzakos A, Gerothanassis IP, Magafa V, Manessi-Zoupa E, Cordopatis P. Comparison of the solution structures of angiotensin I & II. Implication for structure-function relationship. *Eur J Biochem* 2003;270:2163–73. [PubMed: 12752436]
34. Monks SA, Karagianis G, Howlett GJ, Norton RS. Solution structure of human neuropeptide Y. *J Biomol NMR* 1996;8:379–90. [PubMed: 9008359]
35. Worn A, Pluckthun A. Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* 2001;305:989–1010. [PubMed: 11162109]
36. Lippow SM, Wittrup KD, Tidor B. Computational design of antibody-affinity improvement beyond in vivo maturation. *Nat Biotechnol*. 2007
37. Krebs B, Rauchenberger R, Reiffert S, Rothe C, Tesar M, Thomassen E, Cao M, Dreier T, Fischer D, Hoss A, Inge L, Knappik A, Marget M, Pack P, Meng XQ, Schier R, Sohlmann P, Winter J, Wolle J, Kretzschmar T. High-throughput generation and engineering of recombinant human antibodies. *J Immunol Methods* 2001;254:67–84. [PubMed: 11406154]
38. Knappik A, Ge L, Honegger A, Pack P, Fischer M, Wellnhofer G, Hoess A, Wolle J, Pluckthun A, Virnekas B. Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol* 2000;296:57–86. [PubMed: 10656818]
39. Soderlind E, Strandberg L, Jirholt P, Kobayashi N, Alexeiva V, Aberg AM, Nilsson A, Jansson B, Ohlin M, Wingren C, Danielsson L, Carlsson R, Borrebaeck CA. Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol* 2000;18:852–6. [PubMed: 10932154]
40. Hanes J, Jermutus L, Weber-Bornhauser S, Bosshard HR, Pluckthun A. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proc Natl Acad Sci U S A* 1998;95:14130–5. [PubMed: 9826665]
41. Luginbuhl B, Kanyo Z, Jones RM, Fletterick RJ, Prusiner SB, Cohen FE, Williamson RA, Burton DR, Pluckthun A. Directed evolution of an anti-prion protein scFv fragment to an affinity of 1 pM and its structural interpretation. *J Mol Biol* 2006;363:75–97. [PubMed: 16962610]
42. Chames P, Hoogenboom HR, Henderikx P. Selection of antibodies against biotinylated antigens. *Methods Mol Biol* 2002;178:147–57. [PubMed: 11968483]
43. Griswold KE, Kawarasaki Y, Ghoneim N, Benkovic SJ, Iverson BL, Georgiou G. Evolution of highly active enzymes by homology-independent recombination. *Proc Natl Acad Sci U S A* 2005;102:10082–7. [PubMed: 16009931]
44. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 1983;221:709–13. [PubMed: 6879170]
45. Satow Y, Cohen GH, Padlan EA, Davies DR. Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 Å. *J Mol Biol* 1986;190:593–604. [PubMed: 3097327]
46. Stanfield RL, Fieser TM, Lerner RA, Wilson IA. Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. *Science* 1990;248:712–9. [PubMed: 2333521]
47. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* 1987;196:901–17. [PubMed: 3681981]
48. Almagro JC, Hernandez I, Ramirez MC, Vargas-Madrado E. Structural differences between the repertoires of mouse and human germline genes and their evolutionary implications. *Immunogenetics* 1998;47:355–63. [PubMed: 9510553]

Abbreviations

scFv

single chain antibody variable fragment

HV

hypervariable region

CDR	complemetarity determining region
SDR	specificity determining residue
V_H	variable heavy chain
V_L	variable light chain
PCR	polymerase chain reaction
Ab	antibody
FR	framework region
ELISA	enzyme linked immunosorbent assay
HRP	horseradish peroxidase
AT	angiotensin
BSA	bovine serum albumin
PDB	Protein Data Bank

Figure 1. Structure of the Hu:anti-pep and M:anti-pep V_H libraries

(a) Connolly surface⁴⁴ using a 1.7-Å ratio probe of molecular models of Hu:anti-peptide and M:anti-peptide, on the right and left respectively. At the top, side view of the models. The structures in the bottom present the antigen-binding site seen from the antigen perspective. The invariant V_L chains are colored in brown. V_H scaffolds are colored in gray. Within the antigen-binding site, green represents fully randomized positions, *i.e.* those incorporating all 20 amino acids. Positions diversified to Tyr, Asp, Ala, and Ser are in blue. Positions diversified to incorporate the most common amino acids at each position found in an alignment of anti-peptide antibodies in the Kabat database (see Table 1 for details) are in red. The models were built in the Homology module of Insight II (Accelrys) following standard procedures. For Hu:anti-peptide, the coordinates with PDB code 1mcp⁴⁵ and PDB code 2igf⁴⁶ were used as template to build V_L and V_H , respectively. The coordinates with PDB code lig1 (Jeffrey et al., 1993) were used as template for modeling M:anti-peptide. The side-chains of residues variegated in the libraries were modeled and are shown as Ala residues. (b) Sequence of the V_H chains showing positions variegated in the libraries following the same color code as Figure 1a, except for position 94 in Hu: anti-peptide, which is colored in pink. It is not shown in the model, as it is buried in the protein. Numbering and gaps follow Chothia's convention⁴⁷.

Numbering 30

 ..|.|abcdef.....|

A27 CRASQSVSS-----SYLAW

 | | | | | | | | |

B3 CK**SSQSV**LYSSNNKNYLAW

 | | | | | | | | | | | | | | | |

A27md CK**SSQSV**LYSSNNKNYLAW

Figure 2. Sequence of L1 in A27mdJκ1

The HV loop L1 of germline V_L gene B3, with a long canonical structure, was grafted onto germline V_κ gene A27. L1 of both genes A27 and B3 share a key canonical structural determining residue⁴⁸ (V29, underlined in B3). Sequence identities are indicated with vertical bars.

Table 1**M:anti-pep library target diversity of H1, H2**

Positions in bold correspond to the SDRs of 26-10. Positions in italics correspond to the positions of SDRs from 24 anti-peptide antibody co-structures that have side chains which are solvent exposed in 26-10. See materials and methods for the degenerate codes.

Position	Target diversity Prevalence in natural anti-peptide Abs (%)	Codon	Residues encoded/extra residues	Coverage(%)
H29	F-80 L-10 I-9	HTT	FLI	100
<i>H30</i>	T-57 S-29 K-7 N-3 P-1	AVK	TSNK/R	99
<i>H31</i>	S-35 D-23 T-10 E-9 G-8 N-8 R-4 I-1	RVK	SDTEGNNR/KA	99
<i>H32</i>	Y-71 N-8 T-7 F-4 S-4 G-2 A-1 C-1	WHT	FSYTN/I	96
H33	A-21 G-21 W-20 Y-16 T-9 N-6 F-3 I-1 P-1 S-1	DSG	AGWTS/R	74
		WWT	YNFI	26
H35	H-43 N-22 S-21 D-2 E-2 W-2 Y-2 P-1 K-1 L-1	HMT	HNSYP/T	90
	C-1 A-1			
H47	W-99 Y-1	TGG	WY	100
		TAC		
H50	R-16 G-15 T-13 N-12 E-11 V-10 W-8 Y-5 A-4	RNK	RGTNEVAS/IMKD	87
	S-2	TGG	W	8
		TAT	Y	5
H51	I-94 F-3 S-1 V-1	WYT	IFS/T	99
H52	N-22 S-17 R-15 D-13 H-10 Y-10 L-1	MRK	NSRHYW/KQ	74
		TGG		
<i>H52A</i>	P-44 S-14 T-11 L-6 N-6 F-2 Y-1 G-1	HMT	FSLPT/I	91
<i>H53</i>	G-23 Y-21 N-20 A-13 S-12 K-6 D-3 I-1	TAT	Y	21
		RVN	NKSADG/TER	78
<i>H54</i>	N-39 G-30 S-13 Y-7 D-6 T-4	RRT	GSDN	89
<i>H55</i>	G-58 S-13 Y-11 T-5 V-3 A-2 D-2 E-1 H-1 I-1	RBT	ITSVAG	82
	K-1 M-1 F-1	THT	FSY	12
<i>H56</i>	S-20 G-15 Y-14 N-12 D-11 A-10 T-5 V-4 E-4	DMT	SYTNAD	72
	R-3 C-1 Q-1	RKA	VRG/M	22
<i>H57</i>	T-80 P-7 V-3 S-2 A-2 I-1 K-1 L-1 Y-1	NYA	LSPITVA	98
<i>H58</i>	Y-26 N-23 S-13 K-10 T-8 H-7 F-4 D-4 E-2 A-1	THT	FSY	44
		VMN	TAHNKDE/PQ	56

Table 2
Introduction of diversity in the H3 loop of M:anti-pep

Sub-libraries encoding H3 regions of different lengths were constructed by introducing diversity at the positions shown in italics. The codon scheme used at those positions is indicated. NNS encodes all 20 amino acids; DSG encodes W/S/A/G/T/R.

Oligo	CDR-H3 positions 95-102												Length		
	93	94	95	96	97	98	99	100	100a	100b	100c	100d		101	102
NNS-10a	A	SGT	NNS	NNS	NNS	NNS	NNS	Y	A	M			D	Y	10
NNS-10b	A	SGT	NNS	NNS	NNS	NNS	NNS	DSG	A	M			D	Y	10
NNS-12a	A	SGT	NNS	NNS	NNS	NNS	NNS	NNS	NNS	Y	A	M	D	Y	12
NNS-12b	A	SGT	NNS	NNS	NNS	NNS	NNS	NNS	NNS	DSG	A	M	D	Y	12
NNS-8a	A	SGT	NNS	NNS	NNS	Y	A	M					D	Y	8
NNS-8b	A	SGT	NNS	NNS	NNS	DSG	A	M					D	Y	8

Alignment of clones isolated from the Hu:anti-pep library

Alignment of clones isolated from the Hu:anti-pep library

(a)																			
CLONE		CDRH1					CDRH2					CDRH3			H3 length	Frequency			
3-23/J _H 4	29	30	31	32	33	35	50	52	52a	53	54	56	58	94	95	96	97	98	NA
Hu:AT1	.	D	L	S	S	Y	T	A	Y	A	S	A	.	R	R	K	D	I	13/15
Hu:AT2	.	S	T	A	D	A	I	G	S	.	D	Y	S	K	N	G	-	-	2/15

(b)																			
CLONE		CDRH1					CDRH2					CDRH3			H3 length	Frequency			
3-23/J _H 4	29	30	31	32	33	35	50	52	52a	53	54	56	58	94	95	96	97	98	NA
Hu:NPY1	.	D	.	A	D	Y	W	.	A	Y	S	.	D	K	R	S	S	H	6/15
Hu:NPY2	.	A	W	A	D	Y	Y	G	A	Y	A	.	S	K	R	T	T	G	9/15

Table 4
Alignment of clones isolated from the M:anti-pep library

(a) Alignment of the randomized positions in clones isolated after two or three rounds of panning of the M:anti-pep library against AT. The sequence of the scaffold V_H from 26-10 is shown at the top. * denotes stop codon. †Clone M:ATB5 appeared 3/12 and 10/12 times in rounds 2 and 3, respectively. (b) Alignment of the randomized positions in clones isolated after five rounds of panning of the M:anti-pep library against NPY. The dots represent amino acids identical to wild type.

(a)

CLONE	CDRH1					CDRH2												CDRH3												H3 length	Frequency
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58	94	95	96	97	98	99	100	100a	100b	100c	100d			
26-10	F	T	D	F	Y	N	Y	Y	I	S	P	Y	S	G	V	T	G	G	S	S	G	N	K	W	A	M	-	-	10	NA	
Rd. 2																															
M:ATB5	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.	D	.	A	.	.	-	-	10	3/12 (13/24)†	
M:ATC10	.	N	A	.	S	T	.	.	.	.	.	.	.	.	.	.	.	R	I	G	W	G	A	M	-	-	-	-	8	1/12	
M:ATC11	I	S	T	I	F	H	W	A	S	R	Y	E	G	.	D	V	Y	.	N	V	Q	G	T	G	.	.	-	-	10	1/12	
M:ATC12	.	.	.	T	F	S	.	.	.	.	.	.	.	.	.	.	.	.	G	V	T	A	M	-	-	-	-	8	1/12		
M:ATE3	.	N	.	I	S	H	.	G	F	R	N	.	D	S	G	I	Y	.	Q	G	L	L	Q	K	T	G	A	M	12	1/12	
M:ATE5	L	.	N	Y	W	S	.	.	.	.	.	.	.	.	.	.	.	.	G	K	T	A	M	-	-	-	-	8	2/12		
M:ATB4	L	.	A	T	F	H	.	.	.	.	.	.	.	.	.	.	.	.	Q	G	V	T	A	M	-	-	-	-	8	1/12	
M:ATC6	L	.	A	T	W	H	W	T	S	R	Y	.	D	F	G	T	D	.	*	N	.	Y	A	M	-	-	-	-	8	1/12	
Rd. 3																															
M:ATB5	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.	.	G	T	.	D	.	A	.	.	-	-	10	3/12 (13/24)†	
M:ATG11	I	N	A	.	F	H	.	.	.	.	.	.	.	.	.	.	.	.	G	V	T	A	M	-	-	-	-	8	2/12		

(b)

CLONE	CDRH1						CDRH2										CDRH3										H3 length	Frequency		
	29	30	31	32	33	35	47	50	51	52	52a	53	54	55	56	57	58	94	95	96	97	98	99	100	100a	100b			100c	100d
26-10	F	T	D	F	Y	N	Y	Y	I	S	P	Y	S	G	V	T	G	G	S	S	G	N	K	W	A	M	-	-	10	NA
Rd. 5																														
M:NPYA2	I	.	G	I	W	P	W	G	S	Q	.	G	G	V	N	I	Y	R	.	I	P	S	V	L	P	T	A	M	12	1/15
M:NPYA3	.	.	K	T	A	S	.	W	S	W	.	.	D	F	G	V	Y	R	T	G	V	M	S	.	.	-	-	-	10	1/15
M:NPYA6	.	S	T	.	W	Y	W	A	S	.	H	G	D	T	Y	.	S	.	L	E	L	Y	A	M	-	-	-	-	8	1/15
M:NPYA9	L	N	.	Y	T	P	W	G	S	R	S	A	.	I	D	I	S	R	D	H	V	Y	A	M	-	-	-	-	8	1/15
M:NPYD8	I	S	K	S	W	H	W	M	T	R	H	.	G	F	G	L	N	.	V	N	Y	A	E	Y	.	.	-	-	10	1/15
M:NPYE3	.	.	R	S	I	T	W	G	F	.	S	A	D	V	A	V	S	R	L	L	M	S	A	M	-	-	-	-	8	1/15
M:NPYE5	I	.	N	S	I	.	W	T	.	R	H	.	.	F	G	V	Q	R	M	A	Q	V	C	R	.	.	-	-	10	1/15
M:NPYE6	L	R	.	Y	G	Y	W	.	T	Y	H	.	D	F	G	A	H	.	C	Q	A	Y	A	M	-	-	-	-	8	1/15
M:NPYG4	I	.	T	N	I	P	W	T	F	R	Y	S	D	V	Y	I	S	R	P	A	M	W	D	Y	.	.	-	-	10	1/15
M:NPYG12	L	S	A	T	W	H	.	A	F	R	S	.	G	S	.	Y	.	Q	P	P	S	A	V	G	R	A	M	12	1/15	
M:NPYH1	I	.	A	I	A	.	W	R	F	R	Y	T	.	T	Y	V	Y	.	I	K	E	Y	A	M	-	-	-	-	8	1/15
M:NPYH8	I	K	T	Y	W	P	W	W	S	W	H	.	D	S	G	V	Y	.	H	W	F	Q	D	P	M	R	A	M	12	1/15
M:NPYSol1	.	S	T	.	W	S	W	V	S	R	Y	G	D	I	A	A	Y	.	P	G	D	Y	A	M	-	-	-	-	8	2/15
M:NPYSol2	L	K	T	.	T	S	.	V	S	R	S	.	D	.	G	P	Y	.	L	E	W	W	A	M	-	-	-	-	8	1/15

BIACore analysis and expression yields of scFv antibody fragments specific to NPY or AT
Freshly prepared monomeric scAb protein was used to determine the k_a and k_d by surface plasmon resonance on a BIAcore 3000 instrument using biotinylated-peptide immobilized on a SA sensor chip, as described in the materials and methods. The expression yield was calculated from the amount of purified scFv protein as determined by a BCA plate assay.

(a)					
Clone	k_a (1/Ms)	k_d (1/s)	K_D (nM)	χ^2	Expression level (mg/L of culture)
Hu:NPY1	177	7.24×10^{-4}	4100	2.73	4.8
Hu:NPY2	9.06×10^3	4.40×10^{-4}	48.6	0.589	0.3
M:NPYSol1	5.06×10^3	3.81×10^{-4}	75.3	0.731	1.4
M:NPYSol2	1.66×10^4	3.03×10^{-4}	18.2	0.489	0.4

(b)					
Clone	k_a (1/Ms)	k_d (1/s)	K_D (nM)	χ^2	Expression level (mg/L of culture)
Hu:AT1	1.24×10^4	0.0139	1130	1.55	1.6
Hu:AT2	2.08×10^3	0.0112	5400	3.61	1.3
M:ATG11	1.13×10^5	5.94×10^{-3}	52.6	0.44	3.1
M:ATES	1.27×10^3	4.62×10^{-3}	3630	0.364	2.8
M:ATB5	2.56×10^5	3.62×10^{-3}	14.2	0.505	1.5