

A Focused Antibody Library for Improved Hapten Recognition

Helena Persson, Johan Lantto† and Mats Ohlin*

Department of
Immunotechnology
Lund University
P.O. Box 7031, SE-220 07
Lund, Sweden

The topography of the antigen-binding site as well as the number and the positioning of the antigen contact residues are strongly correlated with the size of the antigen with which the antibody interacts. On the basis of these considerations, we have designed a focused scFv repertoire biased for haptens, designated the cavity library. The hapten-specific scFv, FITC8, was used as a scaffold for library construction. FITC8, like other hapten binders, displays a characteristic cavity in its paratope into which the hapten binds. In five of the six complementarity-determining regions, diversity-carrying residues were selected rationally on the basis of a model structure of FITC8 and on known antibody structure–function relationships, resulting in variation of 11 centrally located, cavity-lining residues. L3 was allowed to carry a more complex type of diversity. In addition, length variation was introduced into H2, as longer versions of this loop have been shown to correlate with increased hapten binding. The library was screened, using phage display, against a panel of five different haptens, yielding diverse and highly specific binders to four of the antigens. Parallel selections were performed with a library having diversity spread onto a greater area, including more peripherally located residues. This resulted in the isolation of binders, which, in contrast to the clones selected from the cavity library, were not able to bind to the soluble hapten in the absence of the carrier protein. Thus, we have shown that by focusing diversity to the hotspots of interaction a library with improved hapten-binding ability can be created. The study supports the notion that it is possible to create antibody libraries that are biased for the recognition of antigens of pre-defined size.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: antibody evolution; antibody library; focused diversity; hapten; phage display

*Corresponding author

Introduction

The binding properties of antibodies are determined by the chemistry and shape of their antigen-

binding sites. Six hypervariable loops, supported by a conserved β -sheet framework, come together at the tip of the antibody to form a continuous surface, the antigen-binding site.¹ As the sequence and the length of these loops vary among antibodies, they display different types of surface shapes or topographies.^{2–4} It has been shown that the topography of the paratope correlates with the size of the antigen with which the antibody interacts.^{3,5} One of three different topographies is generally adopted: planar for large antigens (proteins); grooves for medium-sized antigens (peptides and carbohydrates); and cavities for small antigens (haptens). Furthermore, studies have shown differences in the number and location of amino acid residues that play an important role in recognition of antigens of different sizes.^{3,6,7} On average, antibodies use five more residues to contact peptides and proteins than

Present address: J. Lantto, Symphogen A/S, Elektrovej, Building 375, DK-2800 Lyngby, Denmark.

Abbreviations used: BSA, bovine serum albumin; CDR, complementarity-determining region; ELISA, enzyme-linked immunosorbent assay; FITC, fluorescein isothiocyanate or the fluorescein moiety of an FITC-modified molecule; H1, H2 and H3, first, second and third loop of the heavy chain; NIP, 4-hydroxy-3-iodo-5-nitrophenylacetyl/acetate; NP, 4-hydroxy-3-nitrophenylacetyl/acetate; scFv, single chain antibody fragment(s); L1, L2 and L3, first, second and third loop of the light chain.

E-mail address of the corresponding author:
mats.ohlin@immun.lth.se

to recognize haptens. While haptens are recognized mainly by residues located centrally in the antigen-binding site, larger antigens have additional contact residues in more peripheral regions.

It has been speculated that the different binding patterns described above may be used for the construction of libraries biased for antigens of predefined sizes,^{4,7,8} but none has been developed so far. In contrast, the number of large repertoire recombinant antibody libraries from which specific antibody fragments can be selected against virtually any target, for example haptens, peptides, carbohydrates and proteins, are increasing steadily.^{9–15} The size of these libraries relates to the ability to successfully select high-affinity binders towards a particular target.^{11,16} Nevertheless, it is likely that only a relatively small part of a general-purpose repertoire is going to be functional if applied to a particular antigen type, and it is therefore reasonable to believe that a smaller but more functional antibody library would provide the same outcome as a large library. Alternatively, an equally large but more functional library would provide better reagents, possibly reducing the need for affinity maturation.

We postulate that the construction of an antibody library biased for small antigens would allow for more efficient selection of hapten-specific antibodies, which are of major interest in a range of applications. In addition to being an important analytical tool in the biomedical field, for instance in measuring the presence of pharmaceuticals, or a potential aid in drug detection programs, hapten-specific antibodies are needed in the agricultural and food industries; for example, for the detection of antibiotics, food additives, bacterial toxins, pesticides and other environmental contaminants.^{17–19} There are examples of therapeutic applications, such as for neutralizing pharmaceutical overdoses, e.g. the digoxin-specific antibody fragment Digitalis-Antidot BM from Roche. A problem often associated with selecting antibodies (*in vivo* and *in vitro*) to haptens is that the antibodies commonly bind to a combination of the hapten and carrier molecule but not to the free hapten in the absence of the carrier,^{20–22} a feature that would make the selected antibodies ineffective in diagnostic and therapeutic applications. By constructing a library with diversity focused on the residues that commonly interact with the hapten and not with the carrier in hapten-carrier conjugates, we believe that antibodies less prone to bind the carrier can be selected.

In this study, we have explored the above hypothesis by constructing an antibody library biased for haptens. We have designed and created a synthetic antibody library that has been used successfully to find highly specific binders to a range of haptens, independently of the carrier molecule. The library was created by introducing diversity into the residues lining the proposed cavity of a previously described fluorescein-binding antibody fragment.¹⁴ The hapten-binding capability

of this library was compared to that of a general-purpose repertoire. To our knowledge, this is the first report of the creation of an antibody repertoire with a built-in specificity for antigens of pre-defined size.

Results

Design and construction of combinatorial library

As a starting point for library construction the hapten (fluorescein isothiocyanate; FITC)-specific scFv, FITC8, was chosen. FITC8 was originally derived from a synthetic scFv library that had been constructed by shuffling human complementarity-determining region (CDR) sequences of different origins into a single framework consisting of the human *IGHV3-23* and *IGLV1-47* genes.¹⁴ Clones based on this framework, including FITC8, have been shown to be highly stable, well expressed in *Escherichia coli* and displayed easily on phage.^{10,23,24} FITC8, like hapten-binders in general, displays a cavity in its antigen-binding site into which the antigen binds (Figure 1). We hypothesized that by diversifying the FITC8 cavity, a library with the ability to bind a range of different haptens could be created.

Different strategies were used to identify what residues would be suitable for supporting library variability. We started by studying the model structure of FITC8 and included residues that lined the cavity. However, we excluded most

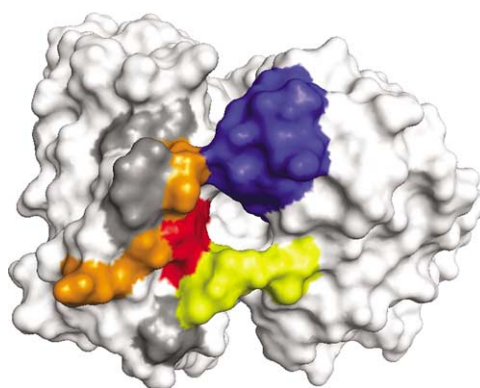


Figure 1. Structure of the library scaffold, FITC8, with the residues targeted for mutation shown in red (H1), orange (H2) and yellow (H3). The full variability of the n-CoDeR library¹⁴ was introduced into L3 (blue). Heavy chain residues N31, Y56 and F59 (grey) were back-mutated to their respective germline sequence. H34 of L1 was also mutated but is not visible in this view. In Table 1, the positions chosen for modifications are listed together with the designed variability. In addition, length variation of H2 was allowed, creating two sub-libraries, Lib1 and Lib2, which together constitutes the cavity library. The Figure was generated with PyMol (<http://pymol.sourceforge.net/>) using Protein Data Bank coordinates obtained from the program MODELER.¹⁴

Table 1. Mutagenesis strategy for the cavity library

Target positions ^a	Designed modifications
H1-N31 (32)	S
H1-W33 (34)	FHLYW
H1-S35 (40)	HNRS
H2-G50 (55)	GRYW
H2-S52 (57)	NSTY
H2-N53 (59)	FINSTY
H2-Y56 (62)	S
H2-Y58 (66)	FHLY
H2-F59 (67)	Y
H3-G95 (107)	GS
H3-D96 (108)	DGNS
H3-G97 (109)	GS
H3-S98 (113)	GS
L1-H34 (40)	HNH

^a Residues are numbered according to Kabat,²⁶ and the corresponding IMGT numbering (<http://imgt.cines.fr>)²⁷ is given in parentheses.

positions that were not considered to be important contact residues in hapten binders, according to analyses done by MacCallum *et al.*³ and Annemarie Honegger *et al.*[†].²⁵ In addition, framework residues were excluded even though these have been shown to be important antigen contacts in hapten binders,^{3,7} as we believed these to be required for structural integrity. As a result, nine amino acid residues, situated in the first, second and third loop of the heavy chain (H1, H2, H3) and the first loop of the light chain (L1) (Figure 1 and Table 1), were chosen for limited substitutions (W33_H, S35_H, G50_H, Y58_H, G95_H, D96_H, G97_H, S98_H, and H34_L) (the Kabat numbering system is used throughout the text if nothing else stated²⁶). In addition, one residue of H1 (N31) and two of H2 (Y56 and F59), situated in the periphery of the antigen-binding site, were mutated back to the residue found in the germline sequence from which the CDR originated, IGHV3-7 and IGHV3-23. We reasoned that these residues, because of their location outside the cavity, might be specific for the carrier protein (streptavidin) that was used originally in the selections to obtain FITC8.¹⁴ No residues were targeted in the second loop of the light chain (L2). It is situated at the periphery of the antigen-binding site and is therefore not expected to make contact with the hapten.

The full variability of the n-CoDeR library¹⁴ was allowed in L3 (Figure 1). Over 70% of the L3s of this library have their origin in IGLV1-36/44, IGLV1-40 and IGLV1-47, the latter being the gene of FITC8. Since these genes show a high degree of sequence similarity, we reasoned that the introduction of any of these L3s into the FITC8 scaffold would most likely lead to sustained cavity topography. The additional diversity that was introduced into this area due to somatic hypermutations and junctional diversity allowed for variability of the sequence length, possibly giving rise to cavities of variable

depth, and provided additional sequence diversity, including variations of positions 91_L and 96_L (IMGT positions 107 and 116, respectively²⁷), which are critical contact residues in hapten recognition.^{3,6,7}

Moreover, since longer H2 loops have been shown to improve the binding of murine hapten-specific antibodies,^{28,29} we decided to introduce length variation of this loop. In addition, the FITC8 framework has been demonstrated to tolerate insertions in H2 with respect to both structure and antigen-binding capacity.⁸ We therefore hypothesized that an insertion in H2 potentially would allow for greater contact with the hapten. Thus, in addition to a library (Lib1) having the original H2 length of FITC8 (17 amino acid residues), a second library version (Lib2) was created, having in addition to the modifications described above, an insertion of two amino acid residues at the tip of the loop, at positions 52b and 52c (between positions 58 and 59 according to the IMGT nomenclature²⁷). In both libraries, the flanking residues (S52 and N53) were mutated, since these might become important contact residues if a conformational change were to occur in the extended H2 loop.

The variability introduced into the library was restricted mainly to what is found in the germline genes that make up the CDR loops of FITC8,¹⁴ together with what is commonly found in hapten binders.[†]²⁵ Degenerate codons were designed for each position, chosen so as to include as many of the intended residues as possible without encoding a large number of extra residues (Figure 6). In H3, small residues were generally chosen to maintain the structure of this loop in upholding the cavity topography. We believed that larger residues, which are commonly found at these positions in hapten binders, would result in blocking of the cavity, thus making it less accessible to hapten binding.

To summarize, in addition to the modified third loop of the light chain (L3), a total of 14 residues in H1, H2, H3 and L1 were targeted for modification (Figure 1 and Table 1). Two library variants (Lib1 and Lib2), together designated the cavity library, were created, differing only in the length of H2. The theoretical size of each of the libraries was estimated to be approximately 7×10^5 , not counting the diversity of L3, as this cannot be estimated properly. Lib1 and Lib2 obtained after transformation contained 2×10^7 and 3×10^7 clones, respectively.

To confirm that the intended diversity had been introduced, clones were picked randomly and sequenced. Altogether, 28 random clones were analyzed from the two libraries. All of these contained the insert, two-thirds were shown to be in-frame and mutations not intended by the primer design, possibly caused by the DNA polymerase, were found in three sequences. The diversity of the libraries was confirmed by identification of most of the designed mutations as well as by an expected number of different L3 variants as defined by the genetic origin of this loop (data not shown). After titration, equal amounts of phages from Lib1 and Lib2 were mixed to create the cavity library.

[†] <http://www.biochem.unizh.ch/antibody/>

Selection of scFv with different hapten specificities

The cavity library was subjected to selections against five haptens: the original antigen FITC as well as testosterone, estradiol, 4-hydroxy-3-iodo-5-nitrophenylacetyl/acetate (NIP) and 4-hydroxy-3-nitrophenylacetyl/acetate (NP). Following two or three rounds of selection, the specificity of individual clones was analyzed by phage enzyme-linked immunosorbent assay (ELISA). As illustrated in Table 2, we were able to efficiently isolate specific binders against four of the five haptens. The proportion of positive clones ranged from 60% to 95%.

Sequencing revealed that only a small fraction of the positive clones originated from Lib2, the library with an insertion of two amino acid residues in the H2 loop (Figure 2). Therefore, to investigate the role of this insertion further, separate selections on NIP-bovine serum albumin (BSA) using only the Lib2 part of the cavity library were performed. Indeed, specific clones could be selected from this library when used without competition from Lib1. After selection, 75% of the clones were specific for the target antigen (Table 2). We conclude that the approach used to create diversity, with respect to amino acid replacements and loop length variation, could be harbored in the FITC8 scaffold and allowed for different specificities to be established.

Binding to the soluble hapten

In order to assess the binding to the soluble hapten, competitive phage ELISA was performed on positive clones from each hapten target. As Lib1 and Lib2 were created on the basis of an FITC-specific antibody sequence, the apparent affinity of the FITC-selected clones was, as expected, higher than that for scFv selected on other targets, having affinities of the same order of magnitude as that of the wild-type (Table 3). Five of the six testosterone-selected clones were inhibited successfully by

testosterone. Analysis of the NIP clones selected from the cavity library revealed a broad range of affinities, with the strongest binders having IC_{50} values in the range 100–1000 ng/ml. Similar results were obtained for the NIP-specific clones originating from Lib2. The estradiol-specific clones were not inhibited by soluble estradiol (Table 2). A possible explanation can be found in the structural differences of the panning antigen, estradiol-6 α -biotin, and estradiol.

Additionally, cross-reactivity studies showed that FITC8, the original clone, had some reactivity against soluble testosterone, although very high concentrations of steroid were required to reach IC_{50} (0.02% cross-reactivity). However, this reactivity was lost among the newly FITC-selected clones of the cavity library. Also, none of the testosterone-selected or NIP-selected clones showed any sign of cross-reactivity against fluorescein (data not shown).

Cavity library versus a conventional antibody library

The behavior of the cavity library, with diversity restricted mainly to a proposed central cavity, was compared to a conventional library, carrying diversity in a considerably larger part of the binding site. For this purpose, we employed a sub-library of n-CoDeR,¹⁴ and compared selections performed on NIP-BSA. As expected, clones specific for the selected antigen could be identified from both libraries (Table 2). When analyzed using competitive phage ELISA, these scFv were blocked efficiently from binding to immobilized antigen by the addition of soluble NIP-BSA conjugate. However, when performing the same kind of analysis using the soluble hapten as inhibitor, none of the clones from the n-CoDeR sub-library had measurable IC_{50} values within the range of concentrations used in this study (Table 3 and Figure 3), indicating a higher degree of carrier dependence of the n-CoDeR-originating clones.

Table 2. Isolation of hapten-specific phage-scFv clones

Library	Library size	Panning antigen	Number of panning rounds	Number of clones specifically binding the panning antigen	Number of unique clones	Number of clones binding to the soluble hapten
Cavity	5×10^7	FITC-b-SA ^a	2-3	18/20	18/18	6/6
Cavity	5×10^7	Testosterone-b-AV ^b	2-3	19/20	19/19	5/6
Cavity	5×10^7	NIP-b-SA	2-3	12/20	12/12	6/6
Cavity	5×10^7	Estradiol-b-AV	2-3	19/20	18/19	0/6
Cavity	5×10^7	NP-b-AV	2-3	0/20	–	–
Lib2	3×10^7	NIP-b-SA	2-3	15/20	12/15	6/6
n-CoDeR sub-library	4×10^8	NIP-b-SA	2-3	8/20	8/8	0/6

After two and three rounds of selection against different target haptens, phage was prepared from individual clones and analyzed for binding specificity by phage ELISA. Clones were considered positive if the signal to the target antigen was at least five times higher than that to control antigens. The number of unique clones was determined by sequencing, and binding to the soluble hapten was assessed by competitive phage-ELISA.

^a b-SA, biotin-streptavidin.

^b b-AV, biotin-avidin.

	H1		H2						H3				L1	L3														Additional substitutions	
Kabat IMGT wt	33 W	35 S	50 G	52 S	52b 58A	52c 58B	53 N	58 Y	95 G	96 D	97 G	98 S	34 40* H	89 A	90 A	91 W	92 D	93 D	94 S	95 L	95a S	95b G	95c 115	96 R	97 V	116	117		
FC-1	.	.	.	N	.	.	I	L	S	N	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	
FC-2	Y	.	Y	.	.	.	T	L	S	N	.	G	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	
FC-4	.	.	.	N	.	.	Y	H	S	N	.	.	.	.	.	.	.	A	.	.	N	A	.	.	.	L	.		
FC-5	F	.	.	T	.	.	T	F	S	G	S	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
FC-6	.	.	.	.	.	.	T	F	S	N	.	.	.	.	T	.	.	.	.	.	.	A	.	.	V	.	.		
FC-7	.	.	.	.	.	.	T	H	S	G	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
FC-9	Y	.	.	T	.	.	I	H	S	N	S	.	.	.	.	.	.	.	.	.	N	.	.	.	.	.	.		
FC-10	H	.	.	T	.	.	I	L	S	S	S	G	.	.	.	.	.	.	.	.	D	.	.	.	.	.	.		
FC-11	L	.	.	.	K	A	Y	H	.	N	.	.	Y	.	.	.	.	.	.	V	N	.	.	.	V	.	.		
FC-12	.	.	.	.	.	.	S	.	.	N	S	.	N	G	T	.	.	.	N	.	N	.	.	.	.	.	M		
FC-13	F	.	.	N	.	.	Y	.	.	N	S	G	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.		
FC-14	.	.	.	.	.	.	F	H	.	S	S	G	.	.	S	.	.	V	.	.	.	.	.	.	.	.	.		
FC-15	.	.	.	N	.	.	.	.	S	N	.	G	.	.	.	.	.	.	.	.	N	.	.	.	.	.	.		
FC-16	Y	.	.	N	.	.	S	.	S	N	.	.	N	.	T	.	.	.	.	.	N	.	.	.	.	.	.		
FC-17	.	.	.	T	.	.	T	H	.	G	.	G	.	S	T	.	.	.	.	.	.	.	.	.	.	.	.		
FC-18	.	.	.	.	.	.	T	H	S	S	.	G	N	Q	S	Y	.	N	.	.	R	.	.	.	.	.	.		
FC-19	.	.	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
FC-20	L	.	.	.	.	.	S	F	.	S	S	.	.	.	T	.	.	.	.	V	N	.	.	.	.	.	.		
EC-1	.	.	.	N	.	.	S	H	.	G	.	G	.	Q	S	Y	.	S	.	V	.	G	P	.	V	.	.		
EC-2	H	N	.	T	.	.	Y	.	.	S	S	.	N	Q	S	Y	.	S	G	.	.	.	.	.	.	.	.		
EC-3	.	N	.	N	.	.	S	.	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-4	H	.	.	T	.	.	F	F	.	.	S	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-6	H	.	.	Y	K	A	Y	.	.	N	S	.	Y	.	.	.	.	.	.	.	N	.	.	.	.	.	.		
EC-7	.	.	.	N	.	.	T	F	S	S	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-8	H	.	.	Y	.	.	T	F	.	.	S	.	Y	.	T	.	.	.	.	V	.	.	.	.	.	.	.		
EC-9	H	.	.	T	K	A	T	H	.	.	S	.	Y	.	S	.	.	.	.	.	.	.	.	.	.	.	.		
EC-10	Y	N	.	.	.	.	T	.	S	S	.	.	.	Q	S	Y	.	N	.	.	.	.	.	.	.	.	.		
EC-11	H	.	.	T	.	.	Y	F	.	G	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-12	H	.	.	Y	K	A	Y	.	.	S	.	G	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-14	H	.	.	T	.	.	S	F	.	N	S	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.		
EC-15	.	.	.	N	.	.	S	H	S	N	.	G	.	Q	S	Y	.	N	T	.	.	.	.	.	.	.	.		
EC-16	H	.	.	T	.	.	S	F	.	.	S	.	Y	.	T	.	.	.	.	.	.	.	.	.	.	.	.		
EC-17	H	.	.	T	.	.	Y	F	.	.	.	.	Y	.	.	.	.	.	N	.	.	A	.	.	.	.	.		
EC-18	.	.	.	N	.	.	Y	F	S	S	.	.	Y	.	S	.	.	.	.	.	.	.	.	.	.	.	.		
EC-19	.	.	.	.	.	.	F	.	.	G	.	G	.	Q	S	Y	.	S	.	.	T	.	.	.	.	.	.		
EC-20	.	.	.	N	.	.	H	.	.	N	S	G	Y	.	.	.	.	.	R	.	N	A	.	.	.	.	.		
TC-1	H	N	W	.	.	.	Y	H	.	N	S	G	Y	.	T	.	.	.	.	.	.	.	S	.	V	.	.		
TC-2	.	.	.	N	.	.	.	.	.	N	S	G	.	.	G	.	.	.	.	.	.	D	.	V	.	L	.		
TC-3	Y	.	.	T	K	A	I	.	.	N	S	.	N	.	T	.	.	.	.	.	N	.	H	.	V	.	.		
TC-5	.	.	.	N	.	.	S	H	.	.	S	G	.	.	.	.	.	.	K	.	N	A	.	.	V	.	.		
TC-6	.	.	.	N	.	.	I	H	.	N	.	G	.	G	T	.	.	.	N	.	S	.	.	.	V	.	.		
TC-7	.	R	.	N	.	.	.	.	.	G	.	.	.	E	T	.	.	T	N	T	H	.	.	.	V	.	.		
TC-8	.	.	.	T	.	.	Y	H	S	S	.	G	Y	.	.	.	.	.	.	.	.	.	.	.	P	.	.		
TC-9	L	.	.	T	.	.	T	H	.	N	S	.	Y	.	.	.	.	.	.	.	N	A	.	.	V	.	.		
TC-10	.	R	.	.	.	.	I	.	S	.	S	.	Y	.	.	.	.	.	.	.	.	.	.	E	.	L	.		
TC-11	F	.	.	Y	.	.	T	.	.	N	S	G	.	.	.	.	.	.	.	M	T	.	.	A	.	L	.		
TC-12	Y	N	W	.	.	.	F	L	S	.	.	.	N	S	S	Y	T	G	.	S	.	.	.	T	.	L	.		
TC-13	Y	.	.	T	.	.	Y	.	.	.	.	.	Y	.	T	.	.	.	.	.	N	.	H	.	V	.	.		
TC-14	Y	.	.	N	.	.	.	.	.	N	.	.	Y	.	.	.	.	.	.	.	K	V	.	.	V	.	.		
TC-15	.	.	W	N	.	.	T	L	.	G	.	.	N	.	.	.	.	.	.	.	N	.	.	.	V	.	.		
TC-16	H	.	W	N	.	.	.	.	.	N	S	.	N	.	T	.	.	.	.	.	N	.	.	.	V	.	.		
TC-17	H	.	.	T	.	.	T	.	S	.	.	G	Y	.	T	.	.	A	.	.	.	.	.	.	P	.	.		
TC-18	Y	.	.	.	K	A	.	L	.	N	.	.	Y	.	T	.	.	H	R	.	L	A	.	G	.	.	.		
TC-19	H	.	.	.	K	A	T	H	S	G	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	P	.	.		
TC-20	.	.	.	N	.	.	Y	.	.	G	S	.	.	.	T	.	.	.	G	.	.	R	.	.	V	.	.		
NC-1	L	R	.	T	.	.	Y	.	.	G	.	.	N	S	S	Y	A	G	.	N	F	.	.	.	W	.	.		
NC-2	F	R	.	N	.	.	F	.	.	G	S	G	N	.	.	.	.	.	.	.	.	.	.	.	W	.	.		
NC-3	.	R	.	T	.	.	S	.	.	N	.	G	N	.	.	.	.	.	.	.	N	.	.	.	W	.	.		
NC-4	.	R	.	.	.	.	I	F	.	N	S	G	N	.	V	.	.	.	.	.	G	.	.	.	W	.	.		
NC-5	.	R	.	N	.	.	Y	.	.	N	S	G	N	.	.	.	.	.	.	.	N	.	.	.	W	.	.		
NC-10	Y	R	.	.	.	.	Y	F	.	N	S	.	.	.	.	.	.	G	.	.	.	.	.	.	W	.	.		
NC-11	.	R	.	N	.	.	T	.	.	N	.	G	N	.	S	.	.	.	.	G	.	.	.	.	W	.	.		
NC-13	Y	R	.	.	.	.	F	.	.	N	.	.	N	.	T	.	.	.	.	.	.	.	.	.	W	.	.		
NC-15	.	R	.	.	.	.	F	F	.	S	S	G	.	.	.	.	.	N	.	.	R	A	.	.	W	.	.		
NC-16	L	R	.	N	.	.	S	F	.	G	S	G	Y	M	I	.	H	N	.	A	.	.	.	W	.	.	.		
NC-19	F	R	.	N	.	.	.	H	.	G	S	G	.	.	.	.	.	K	T	.	F	.	.	.	W	.	.		
NC-20	.	R	.	.	.	.	Y	.	.	S	.	.	.	.	T	.	.	.	.	.	K	.	.	.	I	.	.		
NL2-1	H	R	.	.	K	T	T	.	.	N	S	.	.	S	T	.	.	Y	.	.	A	.	.	P	.	L	.		
NL2-4	.	R	.	.	K	A	.	H	.	N	.	G	.	.	.	.	.	.	.	.	.	.	.	.	W	.	.		
NL2-5	.	R	.	N	K	A	Y	H	.	G	.	G	N	Q	S	Y	.	R	.	.	R	.	.	.	W	.	.		
NL2-8	H	R	.	T	K	A	Y	H	.	G	S	G	Y	.	.	.	.	G	.	.	K	.	.	.	W	.	.		
NL2-9	.	.	.	T	K	A	Y	.	S	G	.	.	N	.	.	.	.	G	.	.	K	.	.	.	W	.	.		
NL2-11	.	R	.	.	K	A	T	.	.	G	.	.	Y	Q	S	Y	.	S	.	.	.	.	P	.	W	.	.		
NL2-12	F	R	.	T	K	A	F	.	.	N	.	.	N	Q	S	Y	.	R	.	.	R	.	.	.	V	.	.		
NL2-14	H	.	.	T	K	A	F	F	S	N	.	.	Y	.	.	.	.	G	R	.	K	.	.	.	W	.	.		
NL2-16	.	R	.	N	K	A	T	.	.	G	S	G	.	Q	S	Y	.	R	.	.	T	.	.	.	W	.	.		
NL2-17	L	R	.	N	K	A	S	F	.	G	S	G	Y	S	S	Y	T	N	T	N	T	H	.	.	W	.	.		
NL2-19	Y	R	.	.	K	A	T	F	.	G	S	.	Y	.	S	.	.	N	.	.	.	.	.	.	W	.	.		
NL2-20	F	R	.	.	K	A	Y	H	.	G	S	G	Y	S	S	.	.	G	.	.	A	.	.	.	W	.	.		

Table 3. Characterization of scFv expressed on phage and in soluble format

	Phage format		Charge at pH 7.4 ^b	scFv format	
	IC ₅₀ (soluble hap- ten) ^a (ng/ml)	IC ₅₀ (NIP-BSA) (ng/ml)		IC ₅₀ (soluble hap- ten) (ng/ml)	K _D (M) ^c
wt	1.5		−1	1.0	3×10 ^{−9}
FC-1	3.2		0	1.2	4×10 ^{−9}
FC-2	2300		0		
FC-6	388		−1		
FC-14	0.4		1	0.3	8×10 ^{−10}
FC-15	0.5		0	0.7	2×10 ^{−9}
TC-1	NR ^d		−1		
TC-2	550		−2		
TC-5	20		−1	110	2×10 ^{−7}
TC-7	1500		0		
TC-13	9600		−2		
TC-20	100		0	110	4×10 ^{−7}
NC-1	4500	70	2		
NC-2	110	280	0	780	3×10 ^{−6}
NC-5	100	420	0	1600	3×10 ^{−6}
NC-11	2600	9600	0		
NC-13	340	1600	0		
NC-20	3000	240	1		
NL2-5	NI ^e	1400	4		
NL2-8	1100	150	3		
NL2-9	430	380	2	1170	1×10 ^{−6}
NL2-11	1500	70	2		
NL2-14	110	40	3	280	9×10 ^{−7}
NL2-20	600	60	2		
NN-1	NR	710	5		
NN-2	NI	200	6	NI	
NN-9	NR	250	3		
NN-13	NI	740	4		
NN-14	NI	200	6		
NN-20	NR	160	6	7900	7×10 ^{−7}

^a IC₅₀-values, defined as the concentration of soluble hapten that blocked 50% of the phage/scFv binding to the immobilized target hapten–protein complex, were determined by competitive ELISA analysis. The IC₅₀ values of the clones correspond to inhibition by their respective target hapten, i.e. the FITC-selected clones (FC) were analyzed for binding to fluorescein, the testosterone-selected (TC) to testosterone and the NIP-selected clones (NC, NL2, NN) to NIP-caproic acid. FC, TC and NC-clones originate from the cavity library, NL2-clones originate from Lib2, while NN-clones were selected from the n-CoDeR sub-library. IC₅₀ values for NIP-BSA were determined also for the NIP-selected, phage-displayed antibody fragments (second column).

^b The total charge of the scFv were estimated using tools available on www.vitalonic.narod.ru/biochem/index_en.html

^c The affinities were determined by surface plasmon resonance. The FC-clones were analyzed on FITC-biotin-streptavidin, the TC-clones on testosterone-biotin-streptavidin and the NIP-specific clones on NIP-BSA.

^d NR indicates that some inhibitory effect was observed, but the IC₅₀ value was not reached within the concentration range used in this study (<10 000 ng/ml).

^e NI indicates that no inhibitory effect was observed.

Characterization of soluble scFv

On the basis of the results of the competitive phage ELISA, the binders of highest affinity for the different hapten targets were chosen for conversion to soluble scFv format and further characterization. For comparisons, we included two NIP-selected clones from the n-CoDeR sub-library. The purified monomeric proteins retained the specificity for their target antigen (Figure 4)

and the cavity library-originating clones could still be inhibited by soluble hapten (Table 3). The concentration of inhibitory hapten was, in several cases, higher than that for the corresponding phage-displayed scFv, most likely due to the higher concentrations of scFv that had to be used in order to obtain detectable signals.

The affinity of the monomeric antibody fragments was determined by BIAcore analysis. All selected clones showed specific and dose-depen-

picked randomly from panning rounds 2 and 3, respectively. Residues identical with the wild-type sequence (FITC8) are shown with dots, while absent residues are indicated by gaps. Residues marked with asterisks are situated in framework regions according to the IMGT nomenclature.²⁷ The mutations N31S, Y56S and F59S are not included in the Figure. Complete gene sequences are available in GenBank, accession numbers DQ250169–DQ250247.

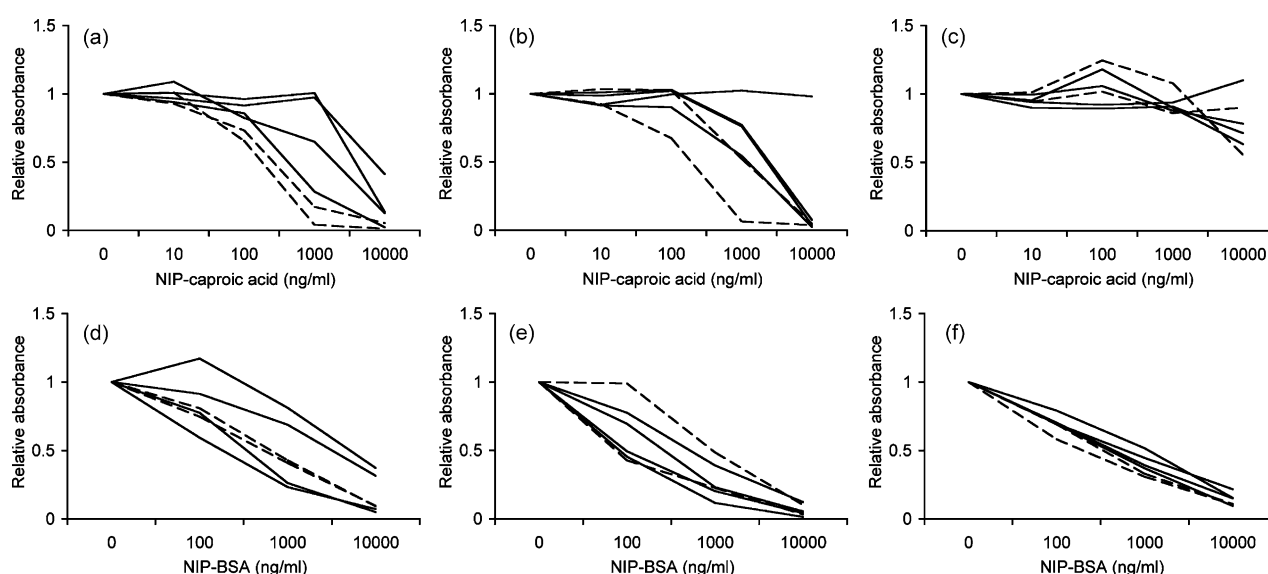


Figure 3. Competitive phage ELISA. Competition for binding was between free hapten (NIP-caproic acid) and hapten-carrier conjugate (NIP-BSA) immobilized on the plate ((a), (b), (c)) and between NIP-BSA in solution and NIP-BSA immobilized on the plate ((d), (e), (f)). NC-clones originate from the cavity library ((a) and (d)), NL2-clones from Lib2 ((b) and (e)) and NN-clones from the n-CoDeR sub-library ((c) and (f)). Six clones were analyzed from each group. The x-axis shows the concentration of the inhibiting antigen (ng/ml) and the y-axis shows the ratio of the signal (A_{490}) in the presence of the inhibiting antigen to that in the absence of soluble antigen. Broken lines indicate clones that were chosen for further analysis.

dent binding to their respective target antigen (Table 3).

Sequence studies

Sequence analysis of the positive clones revealed a large number of unique sequences for the four hapten targets (Table 2). In Figure 2, the sequence of the 11 diversified positions in H1-3 and L1, as well as the whole L3, are given (complete gene sequences are available in GenBank, accession numbers DQ250169–DQ250247). We found that the clones carried a number of key residues that were associated with specific recognition of antigen type. Except for the estradiol-selected and the testosterone-selected clones, which were very similar in sequence, clones selected on the same hapten generally showed a higher degree of similarity compared to clones selected against other antigens.

Sequence comparisons of the NIP-specific proteins originating from the cavity library, Lib2 and n-CoDeR (GenBank accession numbers DQ250248–DQ250255) further revealed a large difference in the estimated net charge of the proteins (Table 3). The clones selected from the cavity library had a charge average around 0.5, and the n-CoDeR clones had an average of 5. The same value for Lib2 was approximately 2.7.

Homology modeling

In order to get a better understanding of the nature of the selected binding sites, we employed an automated homology modeling approach not involving biased intervention by a scientist. All modeled structures, except those with an insertion in H2, showed evidence of a pronounced cavity. In all, this suggested that the selection approach used

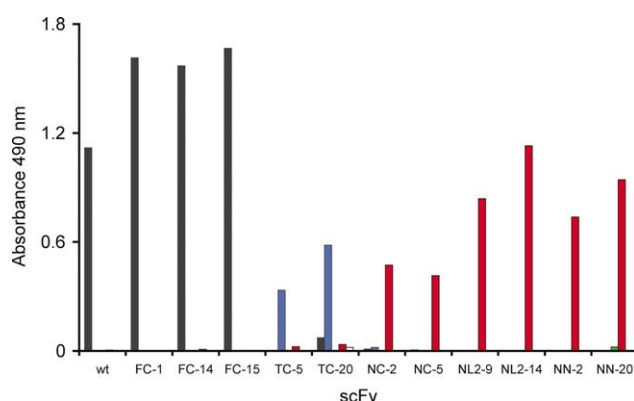


Figure 4. Binding specificity of purified monomeric scFv as determined by ELISA. The antibody fragments were analyzed on microtiter plates coated with FITC-biotin-streptavidin (black bars), testosterone-biotin-streptavidin (blue bars), NIP-BSA (red bars) and on two control antigens, streptavidin (green bars) and BSA (white bars). Numbers are average of duplicates and the relative standard deviation was below 5% for all values.

favorable sequence variants with maintained cavity topography.

Discussion

Several independent studies have identified and characterized differences in binding sites of antibodies that bind antigens of different sizes. Firstly, it has been shown that the topography of the antigen-binding site, which is determined by the length and the sequence composition of the CDR loops, correlates with the nature of the antigen with which it interacts.^{2,3,5} Secondly, differences have been found in the number and the position of residues involved directly in the interaction of the antigen.^{3,6,7} For example, haptens preferentially bind to cavities and are contacted mainly by residues located centrally in the paratope, while proteins that have relatively flat binding surfaces have additional contact residues in more peripheral regions.

In this study, we have explored the possibility of using this correlation between paratope structure and antigen preferences to create a focused antibody library biased for a certain type of antigen, in this case haptens. The hapten-specific antibody FITC8 was used as the structural backbone for library construction. In addition to its fluorescein-binding capacity, FITC8 was shown to bind to testosterone, although the affinity for the latter is much less. This tendency for dual specificity, together with its previously described ability to tolerate loop length insertions,⁸ indicates considerable adaptability of the FITC8 binding-site, a feature we considered to be desirable for our purposes. Its high level of stability and compatibility in *E. coli*-phage systems further suggests that FITC8 is a good candidate for supporting library variability. Diversity-carrying residues were selected rationally on the basis of a model structure of FITC8 and on known antibody structure-function relationships,^{3,25} resulting in variation of mainly centrally located, cavity-lining residues while avoiding diversity in more peripherally located residues. An extension of H2 was introduced into part of the library to allow an evaluation of this factor in the design of a cavity-prone library. In all, five out of six hypervariable loops were designed in this way. Only L3 was allowed to carry a more complex type of diversity, corresponding to that found in the n-CoDeR library.¹⁴

The usefulness of this type of library design was investigated using five different haptens. Phage selections successfully identified a range of specific clones to four of these targets; FITC, estradiol, NIP and testosterone. As the scaffold used for library construction was based on an FITC-specific scFv, the identification of FITC-specific clones can be seen as an evolutionary process of a pre-existing specificity. However, it is interesting to note that these clones lost the original cross-reactivity to testosterone and are specific solely to FITC.

Importantly, the diversification procedure created entirely new binding sites, as evidenced by the complete absence of residual FITC-binding activity in clones selected on testosterone, NIP and estradiol. This was so despite the fact that no measure was taken to select against FITC-binding properties during the panning procedure. Only the smallest hapten, NP, was unable to select for specific binders, instead resulting in an accumulation of non-specific binders. This might suggest that the library design proposed here has a lower limit in the molecular mass of the target either due to geometric constraints or due to the fact that a very small hapten is unable to form a sufficient number of interactions with members of the library. Nevertheless, we conclude that the library design as used here is able to create a library that harbors specific binders for a number of haptens.

In contrast to the NIP, FITC and testosterone-specific clones, the estradiol-specific scFv could not be displaced by its free hapten. A possible explanation can be found in the structural differences of the panning antigen, estradiol-6 α -biotin, and estradiol. A successful isolation of binders depends, especially in the case of haptens, on the structure of the panning antigen. The chemical structure of the linker as well as the coupling position will very much affect the functionality of the hapten molecule. The introduction of the α -amino group at position C6 of the estradiol molecule of the panning antigen has been suggested to change the hapten structure,³⁰ resulting in the selected scFv being non-responsive to the soluble unmodified hapten. Thus, the inability to find clones recognizing soluble hapten in this case is more likely an inherent property of the hapten itself and not of the library *per se*.

Because of the small size of haptens, selection of anti-hapten antibodies (*in vivo* and *in vitro*) requires the hapten to be conjugated to a carrier. As a result, the antibodies and antibody fragments developed against haptens commonly bind to a combination of the hapten and carrier, but not to the free hapten in absence of the carrier.^{20–22} This is based on the fact that in many cases a low molecular mass molecule will not occupy the entire antigen-binding surface of the antibody. The hapten will, as stated above, insert itself into a cavity formed by the antigen-binding site, thus leaving room for additional contacts to be made between the carrier and more peripherally located residues of the binding site. Consequently, the affinity for the isolated hapten will usually be considerably lower than for the hapten-carrier conjugate. An example of this effect is seen clearly when looking at the clones selected from a general-purpose repertoire (as exemplified by a sub-library of n-CoDeR) against NIP-BSA. These clones had similar, if not better, binding affinities to NIP-BSA compared to the cavity library clones as shown by both ELISA and BIAcore analysis. However, there is a dramatic difference in the sensitivity to the free antigen between scFv selected from the two libraries. None of the clones

selected from the library of more conventional design had an IC_{50} value within the range of concentrations used ($<10,000$ ng/ml), while the strongest binder derived from the cavity library had affinities in the 100 ng/ml range, suggesting at least a 100-fold increase in sensitivity to the free hapten of the clones derived from the cavity library. This observation can probably be explained by the difference in positioning of the diversity-carrying residues of the two libraries (Figure 5). The n-CoDeR library has, because of its origin,^{14,31} diversity spread over most residues that comprise the CDR loops, including the centrally situated residues, that are most likely to interact with the hapten, as well as the more peripheral residues, that presumably interact with the carrier protein. The diversity of the cavity library, on the other hand, is focused to the central parts of the antigen-binding site, while diversity is minimal in peripheral areas. The selection procedure of the general-purpose library will therefore enrich for binders targeting a larger surface area compared to binders of the cavity library, which will have little chances of using its peripherally located residues to establish effective interactions with the carrier. Consequently, the clones selected from the cavity library will be directed mainly to the hapten part of the hapten-carrier complex, resulting in better binders to the soluble hapten. Although the carrier-dependence of conventional libraries can be circumvented by alternating between different carriers during the selection procedure or by specifically eluting with the soluble hapten,^{20,21} this effect demonstrates a fundamental difference in functionality between the two types of antibody libraries.

The fact that clones selected from the general-purpose library had additional contact with the carrier was indicated further when comparing the charge of the selected clones from the two repertoires. During panning conditions (pH 7.4), the clones selected from the cavity had an average charge of approximately +0.5, while the same

value for the clones selected from the general-purpose library was +5, much above the average charge of members of this library. Interestingly, when looking at model structures of the latter it turned out that the positive charges were situated mainly in the peripheral part of the paratope, suggesting that these residues interact with the abundant negative charges on the carrier (BSA) rather than with the hapten. In conclusion, the results of the selections from the two libraries demonstrate clearly that the cavity library, with diversity focused to the central area of the paratope, is better suited for binding haptens, in the absence of the carrier, than a general-purpose repertoire.

Sequencing of the clones selected from the cavity library revealed a large number of unique sequences from each hapten target. Most clones of the same specificity could be grouped on the basis of common sequence patterns, and this specificity could be coupled to certain key residues, in particular residue 35 of H1 and residue 96 of L3. Situated close to the bottom of the binding pocket, these are highly conserved within each group of hapten, although differing between them, possibly suggesting that these residues are important specificity-determining residues. Residue 34_L is also highly restricted, although not to the same extent. The importance of these residues as contact residues has been demonstrated in other hapten-binding antibodies.^{3,7} 91_L has also been identified as an important antigen contact. Indeed, the diversity in this position is highly conserved among the selected clones, but does not differ significantly from randomly picked clones of the unselected library (data not shown), making it difficult to draw any conclusion about the function of this residue. The high frequency of glycine in position 50_H of the selected clones compared to the unselected cavity-library (94% and 25%, respectively) indicates this residue is critical in obtaining a hapten-binding pocket. It seems that the large difference in size between glycine and the other residue alternatives

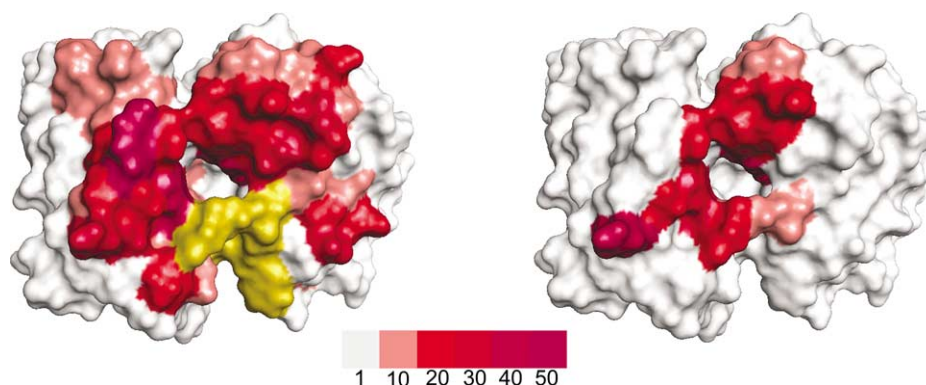


Figure 5. The level of diversity in clones from the n-CoDeR library (left) and the cavity library (right) superimposed on the structure model of FITC8. The H3 loop (residues 95–102) of the n-CoDeR library is shown in yellow. Other residues are color-coded from white to deep red to indicate increasing variability (1–50) as defined by Wu and Kabat.³⁹ The calculated diversity of n-CoDeR is based on 97 random sequences, while the calculations of the cavity library were based on the theoretical variability as defined by the library design.

(R, W, Y), which are found frequently in many hapten-specific antibodies of other origins,²⁵ prevented the incorporation of the latter in functional proteins in this scaffold. We suggest that the diversity restriction in this position is likely due to structural requirements in maintaining the cavity topography of the binding site rather than to contact the antigen. The fact that the scFv genes cluster based on antigen specificity (data not shown) may suggest that these bind their target antigen in the same conformation. A combination of the restricted diversity of the library and the small size of the target molecule may restrict the number of possible epitopes that can be recognized, resulting in very homogenous populations. It is further noteworthy that clones selected on testosterone and estradiol are very similar, which is probably due to the small structural differences that exist between these two steroid molecules. Despite the similarity at the sequencing level, no cross-reactivity of these clones was observed (data not shown).

Since small antigens are contacted mainly by residues located centrally in the antigen-binding site, haptens commonly have very few contacts in the H2 loop, which is generally situated in the periphery of the antigen-binding site.⁷ This is consistent with the proposed structure of the FITC8 scFv. In some antibodies, long H2 together with long L1 have been suggested to promote the formation of cavities and thus increase the hapten-binding ability.² Indeed, it has been shown that an increase in the length of H2 can improve the binding affinity of existing hapten-binding antibodies,^{28,29} most likely due to a conformational change resulting in the relocation of the H2-loop to a more central position of the paratope. As outlined above, we therefore designed a part of the cavity library to exploit this aspect of the binding site to potentially enhance the hapten-binding capacity. When selecting specific variants from a phage pool containing equal numbers of clones with (Lib2) or without (Lib1) an extended H2 loop, the shorter loop tended to be selected preferentially, independent of the antigen used for selection. The fact that it was possible to select specific binders against NIP from the extended H2-loop library when used alone, without competition from Lib1, suggested that an insertion indeed could be tolerated in the FITC8 scaffold to produce functional hapten binders. It was shown that clones from this pool had apparent binding affinities to the soluble hapten similar to those selected from the library lacking the extended H2-loop. However, binders from the Lib2 library tended to have somewhat higher binding affinities for the hapten-carrier conjugate as opposed to the free hapten, possibly suggesting that the insertion increases the contact to the carrier rather than to the hapten. This conclusion is consistent with that proposed by Parhami-Seren and co-workers.²⁹ Sequence analysis of the selected clones further revealed an increase in net charge of the Lib2-clones compared to the clones of Lib1, which cannot be explained entirely by the

introduction of lysine in position 52b. As in the case of the clones selected from the n-CoDeR sub-library, as mentioned above, we believe that these extra positive charges interact with the negative charges on the carrier, thus explaining the stronger binding to the hapten-carrier conjugate of the clones carrying the insertion. The selective disadvantage of Lib2-originating clones in the presence of Lib1, despite similar or higher affinity for the panning antigen and similar levels of display (data not shown), remains unclear. A possible explanation could be that the bacteria encoding proteins carrying the insertion might have a growth disadvantage and are thus selected against during the iterative growth and selection process. In conclusion, we have shown that both sub-libraries used in this study are functional with respect to their capacity to establish hapten-specific binding sites, but the insertion of an extended H2 loop did not provide any additional beneficial effects with respect to hapten recognition.

The affinities of the selected clones are, with the exception of the FITC-specific clones, relatively moderate. However, we believe that the affinities obtained can be improved. It is likely that the standardized procedure used in these studies has not utilized the capacity of the library to its fullest. The high level of diversity of the selected clones suggests strongly that an increased selection pressure can be applied in order to retrieve superior binders; for example, by reducing the amount of selecting antigen or by additional rounds of selection. For structural reasons, with the exception of L3, we did not include buried or framework residues in the library design. Since some of these are known, through indirect effects, to modulate affinity,^{32–34} and targeting of these would most likely lead to further improvements. We are also quite confident that future structural studies of FITC8 together with insights obtained from the current study will help in optimizing the diversity design.

To conclude, by considering the specific binding characteristics of antibodies we have created a focused antibody library that has been used successfully to find highly specific binders to a panel of different haptens, independent of the carrier molecule. Thus, the creation of small libraries intended for the establishment of specific binders of a given nature has been established. We envisage that other types of library designs can be established to target common features of other sets of antigenic targets.

Materials and Methods

Construction of libraries and preparation of phage stocks

The template gene, encoding the anti-FITC scFv (FITC8), used for library construction has been described.¹⁴ In order to create unique restriction sites for the subsequent cleavage reactions, a redundant *EagI*

The libraries were infected by VCSM13 helper phage (Stratagene, La Jolla, CA, USA) as described.³⁶ Phage were harvested by precipitation with polyethylene glycol/NaCl and titrated. Equal amounts of phage from the two libraries were mixed (from hereon designated the cavity library) and subsequently used for selection. For comparison studies, a phage stock of a fraction (approximately 4×10^8 clones) of the n-CoDeR-library¹⁴ was

Phage stocks of random clones from the second and third round of selection were prepared and analyzed for binding specificity by phage ELISA. To ensure specific binding, the clones were tested on the biotinylated target hapten (bound *via* streptavidin or avidin) as well as on

Name	Sequence	Location
a-5'	5'-GGCCGATTTCATTAATGCAGC	Vector specific primer
a-3'	5'-TACTAAAGGTGAATCCAGAGG	Codons H31-H24
b1-5'	5'-CCTCTGGATTACCTTTAGTAGCTATTGGATGMRCTGGGTCCGCCAGGCTCCAGG ^a	Codons H24-H42
b2-5'	5'-CCTCTGGATTACCTTTAGTAGCTATYWCATGMRCTGGGTCCGCCAGGCTCCAGG ^a	Codons H24-H42
b1-3'	5'-TCCTTTCACGGAGTCTGCGTAGWRTGTGCTACCACCGDWACCACAKWAATATATGAGACCCACTCCAGTCCCTT ^a	Codons H65-H43
b2-3'	5'-TCCTTTCACGGAGTCTGCGTAGWRTGTGCTACCACCGDWACCACAKWAATCCHTGAGACCCACTCCAGTCCCTT ^a	Codons H65-H43
b3-3'	5'-TCCTTTCACGGAGTCTGCGTAGWRTGTGCTACCACCGDWCGCTTTACCAKWAATATATGAGACCCACTCCAGTCCCTT ^a	Codons H65-H43
b4-3'	5'-TCCTTTCACGGAGTCTGCGTAGWRTGTGCTACCACCGDWCGCTTTACCAKWAATCCHTGAGACCCACTCCAGTCCCTT ^a	Codons H65-H43
c-5'	5'-CTACGCACACTCCGTGAAGGACCG	Codons H58-H66
c-3'	5'-TTGGCCCCAGAGGACCAGCCACYGCYAYYACYTCCCCGCACAGTAATACACAGC ^a	Codons H105-H88
d-5'	5'-GGCTGGTCTCTTGGGGCCAA	Codons H99-H105
d-3'	5'-CTGGGAGCTGCTGTATACCAAGTDTACATCATAACCTGCCCCGAT ^a	Codons L41-L27c
e-5'	5'-ACTGGTATCAGCAGCTCCCG	Codons L34-L41
e-3'	5'-CTTGGAGCCAGAGAATCGGTC	Codons L66-L60
f-5'	5'-GACCGATTCTCTGGCTCCAAG	Codons L60-L66
f-3'	5'-GCCCTTACGCGTCAGACTGTAG	Vector specific primer

Figure 6. Oligonucleotide primers used for construction of the cavity library, which is composed of two sub-libraries, Lib1 and Lib2, differing only in the length of the H2 loop and consisting of six fragments each (a, b, c, d, e, and f). Primers b1-5' and b2-5' as well as b1-3' and b2-3' are two versions of the same primer introducing different mutations in position 33 of H1 and position 50 of H2, respectively. Primers b3-3' and b4-3' are identical with b1-3' and b2-3' but with the additional insertion of two amino acid residues in H2. A combination of b1-5' and b2-5' together with b1-3' and b2-3' or b3-3' and b4-3' was used for construction of fragment b, with the shorter or the longer H2, respectively. FITC8 was used as template for the construction of all the fragments except for f, which encodes the L3 region. Instead, a plasmid preparation of the n-CoDeR library¹⁴ was used. The fragments were assembled step-wise into scFv genes by overlap-extension PCR. The primers anneal to one or more nucleotides in the given codons.

empty streptavidin or avidin, BSA, and at least one other hapten. Bound phage were detected by the horse-radish peroxidase-conjugated anti-M13 monoclonal antibody (Amersham Pharmacia Biotech Inc., Piscataway, NJ, USA) using *o*-phenylenediamine as chromogen. The absorbance at 490 nm was recorded. Clones were considered positive if the signal to the target antigen ratio was at least five times higher than the signal to the non-relevant antigens ratio. Positive clones were subjected to DNA sequencing (MWG Biotech, Ebersberg, Germany).

Competitive phage ELISA was performed to investigate the binding of the selected scFv to soluble hapten. A fixed concentration of phage, diluted in blocking buffer (0.1% (w/v) gelatin and 0.05% (v/v) Tween 20 in phosphate-buffered saline) to yield signals in the linear range, were pre-incubated with serial dilutions (0.1–10,000 ng/ml) of the free non-conjugated form of the target hapten (fluorescein, NIP-caproic acid, estradiol, or testosterone) for 2 h at 37 °C. The mixture was transferred to plates coated with biotinylated target antigens (bound *via* streptavidin or avidin) and incubated for another 30 min. The detection was performed as described above. IC₅₀ values, defined as the concentration of soluble hapten that blocked 50% of the phage binding to the immobilized hapten–protein complex, were determined as an estimation of the binding affinities of the selected antibody fragments. In order to assess the relevance of the carrier protein for binding, the clones that had been selected on NIP were also inhibited by NIP-BSA (Biosearch Technologies) in the same way as for the soluble hapten.

To confirm binding specificity of the phage-displayed antibody fragments, ELISA analyses were performed on soluble scFv, essentially as described above but with the following changes: scFv fragments were detected using anti-FLAG M2 antibody (Sigma Aldrich) or Tetra-His antibody (Qiagen, Hilden, Germany), together with a rabbit anti-mouse immunoglobulin/horse-radish peroxidase conjugate (DAKO A/S, Glostrup, Denmark).

Production and purification of soluble scFv

Phagemid DNA from clones having the lowest IC₅₀ values in the competitive phage ELISA was isolated and the gene III sequence removed by digestion with *EagI* to allow for production of soluble scFv. The religated plasmids were transformed into *E. coli* and culture of single colonies grown in Terrific Broth containing ampicillin (100 µg/ml). The expression of proteins was induced at A₄₉₀=0.9 by the addition of 1 mM isopropyl thiogalactoside (IPTG) for 16 h. The proteins were purified from the periplasmic fraction by resuspending the bacterial pellet in 1 mg/ml of lysozyme (Sigma-Aldrich) in ice-cold sucrose solution (580 mM sucrose, 30 mM Tris, 1 mM EDTA). After incubation for 1 h at 4 °C, the samples were centrifuged at 48,000g for 15 min and the supernatants, containing the scFv, were filter-sterilized and purified on His Trap HP nickel columns (Amersham Biosciences, Uppsala, Sweden). Finally, the monomers were separated from dimers and higher aggregates by applying the samples to a Superdex 75 preparation grade XK16/70 column (Amersham Biosciences). Protein concentrations were determined spectrophotometrically using extinction coefficients estimated individually for each clone.³⁷

Affinity measurements

The affinity of the purified monomeric antibody fragments was determined by surface plasmon resonance analysis using a BIAcore 3000 instrument (BIAcore AB, Uppsala, Sweden). Proteins were coupled covalently to CM5 sensor chips using amine coupling according to the manufacturer's instructions. Biotinylated testosterone and FITC were captured on the chip surface, to which streptavidin had been coupled, while NIP-BSA was attached directly *via* its free amino groups. The binding of the scFv were measured at different concentrations at a continuous flow-rate of 20 µl/min. All experiments were performed at 25 °C and the chip was regenerated with a pulse of 20 mM HCl, 150 ml NaCl. Response curves were obtained by subtracting the signal of a reference surface from that observed. Streptavidin was used as reference for the FITC and testosterone-coated surfaces, whereas BSA was used for NIP-BSA. The response was analyzed using BIAevaluation 3.0 (BIAcore). Reaction rate kinetics of the FITC-specific clones was calculated using the pre-defined 1:1 Langmuir binding model. The K_D values of the testosterone-specific and NIP-specific clones were determined by plotting steady state (R_{eq}) levels against various concentrations of scFv, followed by analysis using the steady-state affinity model.

Modeling

Homology models of the scFv were generated using the web-based antibody-modeling program WAM†.³⁸ For all but one clone, the accessibility profile method was used as final screen together with the CONGEN iterative algorithm as side-chain building method. NC-2, which according to WAM has a non-kinked H3 loop, was modeled using the VFF model.

Acknowledgements

We thank Dr Roland Carlsson for helpful discussions. This study was supported by grants from the Swedish Research Council and Bioinvent International AB (Lund, Sweden).

References

1. Padlan, E. A. (1994). Anatomy of the antibody molecule. *Mol. Immunol.* **31**, 169–217.
2. Vargas-Madrado, E., Lara-Ochoa, F. & Almagro, J. C. (1995). Canonical structure repertoire of the antigen-binding site of immunoglobulins suggests strong geometrical restrictions associated to the mechanism of immune recognition. *J. Mol. Biol.* **254**, 497–504.
3. MacCallum, R. M., Martin, A. C. & Thornton, J. M. (1996). Antibody-antigen interactions: contact analysis and binding site topography. *J. Mol. Biol.* **262**, 732–745.
4. Collis, A. V., Brouwer, A. P. & Martin, A. C. (2003). Analysis of the antigen combining site: correlations

† <http://antibody.bath.ac.uk/>

- between length and sequence composition of the hypervariable loops and the nature of the antigen. *J. Mol. Biol.* **325**, 337–354.
5. Wilson, I. A., Rini, J. M., Fremont, D. H., Fieser, G. G. & Stura, E. A. (1991). X-ray crystallographic analysis of free and antigen-complexed Fab fragments to investigate structural basis of immune recognition. *Methods Enzymol.* **203**, 153–176.
6. Padlan, E. A., Abergel, C. & Tipper, J. P. (1995). Identification of specificity-determining residues in antibodies. *FASEB J.* **9**, 133–139.
7. Almagro, J. C. (2004). 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. Recogn.* **17**, 132–143.
8. Lantto, J. & Ohlin, M. (2002). Functional consequences of insertions and deletions in the complementarity-determining regions of human antibodies. *J. Biol. Chem.* **277**, 45108–45114.
9. Hoogenboom, H. R. & Winter, G. (1992). By-passing immunisation. Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro. *J. Mol. Biol.* **227**, 381–388.
10. Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L. *et al.* (1994). Isolation of high affinity human antibodies directly from large synthetic repertoires. *EMBO J.* **13**, 3245–3260.
11. Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C. *et al.* (1996). Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nature Biotechnol.* **14**, 309–314.
12. Pini, A., Viti, F., Santucci, A., Carnemolla, B., Zardi, L., Neri, P. & Neri, D. (1998). Design and use of a phage display library. Human antibodies with subnanomolar affinity against a marker of angiogenesis eluted from a two-dimensional gel. *J. Biol. Chem.* **273**, 21769–21776.
13. de Haard, H. J., van Neer, N., Reurs, A., Hufton, S. E., Roovers, R. C., Henderikx, P. *et al.* (1999). A large non-immunized human Fab fragment phage library that permits rapid isolation and kinetic analysis of high affinity antibodies. *J. Biol. Chem.* **274**, 18218–18230.
14. Söderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Åberg, A. M. *et al.* (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nature Biotechnol.* **18**, 852–856.
15. Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellenhofer, G. *et al.* (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57–86.
16. Little, M., Welschhof, M., Braunagel, M., Hermes, I., Christ, C., Keller, A. *et al.* (1999). Generation of a large complex antibody library from multiple donors. *J. Immunol. Methods*, **231**, 3–9.
17. Harris, B. (1999). Exploiting antibody-based technologies to manage environmental pollution. *Trends Biotechnol.* **17**, 290–296.
18. Kramer, K. & Hock, B. (2003). Recombinant antibodies for environmental analysis. *Anal. Bioanal. Chem.* **377**, 417–426.
19. Moghaddam, A., Borgen, T., Stacy, J., Kausmally, L., Simonsen, B., Marvik, O. J. *et al.* (2003). Identification of scFv antibody fragments that specifically recognise the heroin metabolite 6-monoacetylmorphine but not morphine. *J. Immunol. Methods*, **280**, 139–155.
20. Moghaddam, A., Lobersli, I., Gebhardt, K., Braunagel, M. & Marvik, O. J. (2001). Selection and characterisation of recombinant single-chain antibodies to the hapten Aflatoxin-B1 from naive recombinant antibody libraries. *J. Immunol. Methods*, **254**, 169–181.
21. Charlton, K., Harris, W. J. & Porter, A. J. (2001). The isolation of super-sensitive anti-hapten antibodies from combinatorial antibody libraries derived from sheep. *Biosens. Bioelectron.* **16**, 639–646.
22. Kramer, K. (2002). Synthesis of a group-selective antibody library against haptens. *J. Immunol. Methods*, **266**, 209–220.
23. Ewert, S., Honegger, A. & Plückthun, A. (2003). Structure-based improvement of the biophysical properties of immunoglobulin VH domains with a generalizable approach. *Biochemistry*, **42**, 1517–1528.
24. Steinhauer, C., Wingren, C., Hager, A. C. & Borrebaeck, C. A. K. (2002). Single framework recombinant antibody fragments designed for protein chip applications. *Biotechniques*, **Suppl.**, 38–45.
25. Honegger, A. & Plückthun, A. (2001). Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *J. Mol. Biol.* **309**, 657–670.
26. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, vol. 5, NIH Publication no. 91-3242 U.S. Department of Health and Human Services.
27. Lefranc, M. P., Pommie, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L. *et al.* (2003). IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77.
28. Lamminmäki, U., Pauperio, S., Westerlund-Karlsson, A., Karvinen, J., Virtanen, P. L., Lövgren, T. & Saviranta, P. (1999). Expanding the conformational diversity by random insertions to CDRH2 results in improved anti-estradiol antibodies. *J. Mol. Biol.* **291**, 589–602.
29. Parhami-Seren, B., Viswanathan, M. & Margolies, M. N. (2002). Selection of high affinity p-azophenyarsenate Fabs from heavy-chain CDR2 insertion libraries. *J. Immunol. Methods*, **259**, 43–53.
30. Lacorn, M., Fleischer, K., Willig, S., Gremmel, S., Steinhart, H. & Claus, R. (2005). Use of biotinylated 17 β -estradiol in enzyme-immunoassay development: spacer length and chemical structure of the bridge are the main determinants in simultaneous streptavidin-antibody binding. *J. Immunol. Methods*, **297**, 225–236.
31. Borrebaeck, C. A. K. & Ohlin, M. (2002). Antibody evolution beyond Nature. *Nature Biotechnol.* **20**, 1189–1190.
32. Tomlinson, I. M., Walter, G., Jones, P. T., Dear, P. H., Sonnhämmer, E. L. & Winter, G. (1996). The imprint of somatic hypermutation on the repertoire of human germline V genes. *J. Mol. Biol.* **256**, 813–817.
33. Ramirez-Benitez, M. C. & Almagro, J. C. (2001). Analysis of antibodies of known structure suggests a lack of correspondence between the residues in contact with the antigen and those modified by somatic hypermutation. *Proteins: Struct. Funct. Genet.* **45**, 199–206.
34. Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. & Stevens, R. C. (1997). Structural

- insights into the evolution of an antibody combining site. *Science*, **276**, 1665–1669.
35. Engberg, J., Andersen, P. S., Nielsen, L. K., Dziegiel, M., Johansen, L. K. & Albrechtsen, B. (1995). Phage-display libraries of murine and human antibody Fab fragments. *Methods Mol. Biol.* **51**, 355–376.
36. Cicortas Gunnarsson, L., Nordberg Karlsson, E., Albrekt, A. S., Andersson, M., Holst, O. & Ohlin, M. (2004). A carbohydrate binding module as a diversity-carrying scaffold. *Protein Eng. Des. Sel.* **17**, 213–221.
37. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.
38. Whitelegg, N. R. & Rees, A. R. (2000). WAM: an improved algorithm for modelling antibodies on the WEB. *Protein Eng.* **13**, 819–824.
39. Wu, T. T. & Kabat, E. A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Expt. Med.* **132**, 211–250.

Edited by J. Karn

(Received 21 October 2005; received in revised form 22 December 2005; accepted 3 January 2006)
Available online 19 January 2006