Investigation of antigen–antibody interactions using a soluble, non-support-bound synthetic decapeptide library composed of four trillion (4×10^{12}) sequences

Clemencia PINILLA, Jon R. APPEL and Richard A. HOUGHTEN*
Torrey Pines Institute for Molecular Studies, 3550 General Atomics Court, San Diego, CA 92121, U.S.A.

A decapeptide positional-scanning synthetic-peptide combinatorial library (PS-SPCL) made up of four trillion (4×10^{12}) decapeptides was synthesized; its use is illustrated here for the study of a peptide–antibody interaction. This library was prepared by a chemical-mixture approach using a specific ratio of amino acids empirically determined to give approximately equimolar incorporation of each amino acid during each coupling step. Despite the immense number of decapeptides making up each peptide mixture [approx. 200 billion (2×10^{11})], specific sequences having nanomolar affinities for a peptide–antibody interaction could be readily identified. Upon screening this decapeptide PS-SPCL in this well characterized system, the known six-residue antigenic-determinant sequence was found, with the most specific residues appearing to 'walk through' the ten positions of the peptide library. More importantly, it appears that antibody recognition in this system is stronger when the antigenic determinant is located at the C-terminus of the decapeptide library. Individual decapeptides corresponding to sequences derived from the most active peptide mixtures at each position were synthesized to confirm the results of the screening; 15 peptides were found to have IC_{50} values between 0.6 and 9.5 nM, four of which were found to be 5–10 times more active than the known six- and 13-residue control peptides. These results further illustrate the power of the positional-scanning peptide library concept, and extend its practical range to a decamer library composed of four trillion (4×10^{12}) decapeptides.

INTRODUCTION

Current methods that use millions of peptides in various formats to identify individual peptides of diagnostic and therapeutic interest (Geysen et al., 1986; Cwirla et al., 1990; Devlin et al., 1990; Houghten et al., 1991; Lam et al., 1991; Cull et al., 1992; Lowman et al., 1992; Pinilla et al., 1992; Dooley and Houghten, 1993) rely upon the inherent selectivity of biological receptor systems. Unlike the majority of these approaches, synthetic peptide combinatorial libraries (SPCLs) (Houghten et al., 1991, 1992), which use an iterative selection and enhancement process to define the most active sequence, are composed of soluble, non-support-bound hexapeptides that can be used in virtually any existing bioassay. SPCLs have been successfully used for the identification of antigenic determinants (Appel et al., 1992; Houghten et al., 1991, 1992; Pinilla et al., 1993a), novel peptide ligands for the opioid receptor (Dooley et al., 1993; Houghten and Dooley, 1993), enzyme inhibitors (Eichler and Houghten, 1993), and highly potent antibacterial peptides (Houghten et al., 1991, 1992; Blondelle and Houghten, 1994).

A conceptually different approach, which we have termed a positional-scanning SPCL (PS-SPCL), substantially shortens the timeframe required for the identification of high-affinity peptide sequences. In earlier work, the use of hexapeptide PS-SPCLs has enabled the identification of highly specific and active sequences in a single assay involving antigen–antibody (Pinilla et al., 1992) and receptor–ligand (Dooley and Houghten, 1993) interactions. In order to test the limits of the PS-SPCL approach, we have synthesized an L-amino acid decapeptide library and use it in the study of a peptide–antibody interaction. The monoclonal antibody (mAb 17/9) used in the current study was raised against a 36-residue peptide from the haemagglutinin of influenza virus (HA1 75–110), and was found to recognize a 13-residue region using overlapping peptides (Wilson et al., 1984). The specificity and relative positional importance at the amino acid level of the antigenic determinant -DVPDYA- of this peptide–antibody interaction have been well characterized using a complete set of substitution analogues (Pinilla et al., 1993b). It was found that the two aspartic acid residues and the tyrosine residues at positions 1, 4 and 5 respectively were virtually irreplaceable if antibody binding was to be maintained. Conversely, the valine, proline and alanine residues at positions 2, 3 and 6 respectively were found to be replaceable to various extents by other amino acids. This prior knowledge of the specificity of this peptide–antibody interaction at the amino acid level facilitates the analysis of the results obtained upon screening the enormously diverse peptide library described in the present study. The screening of this decapeptide PS-SPCL against a known interaction also serves as a foundation for its use in other systems.

EXPERIMENTAL

Peptides synthesis

The ten positional peptide libraries of the decapeptide PS-SPCL, as well as subsequent peptide mixtures and individual peptides, were synthesized using simultaneous multiple peptide synthesis methodology (Houghten, 1985; Houghten et al., 1986), methylbenzhydrylamine polystyrene resin, and t-Boc protected L-amino acids. Peptide mixture resins were prepared using a predetermined ratio of 18 of the 20 natural amino acids coupled as a mixture (Dooley and Houghten, 1993). Mixtures used as reference

Abbreviations used: mAb, monoclonal antibody; PS-SPCL, positional-scanning synthetic-peptide combinatorial library; the terms 'billion' and 'trillion' retain their 'U.S.' meanings of 10^9 and 10^{12} respectively.

* To whom correspondence should be sent.

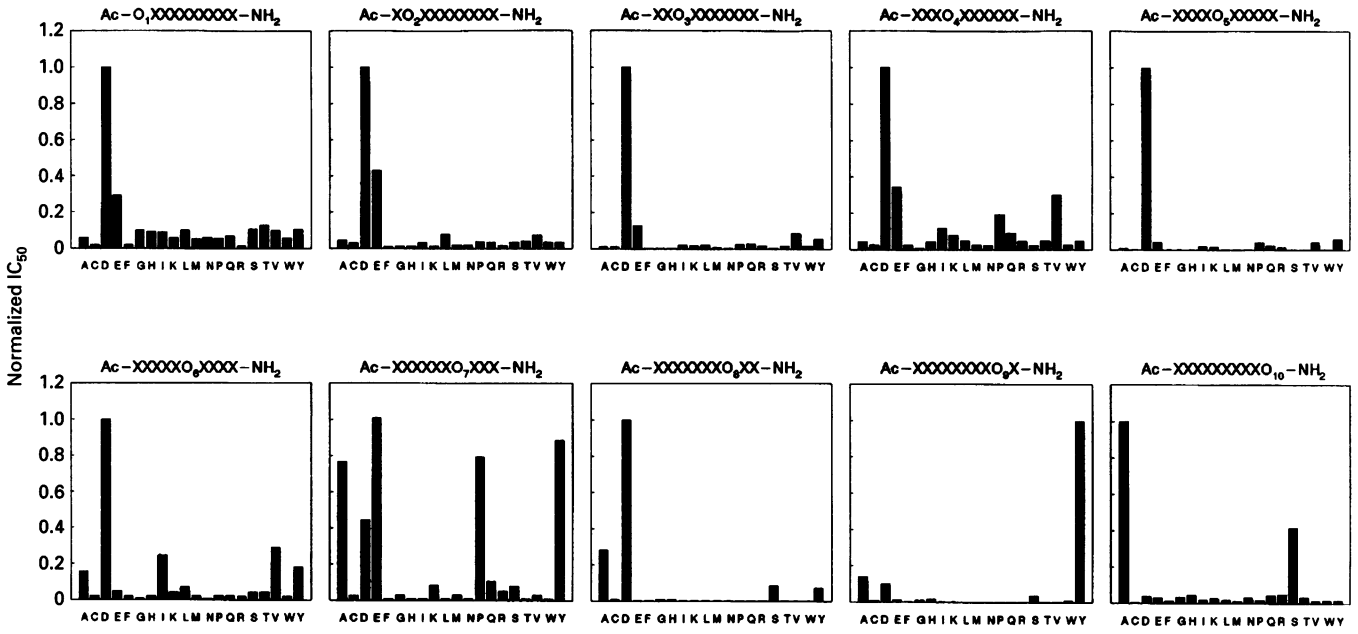

Figure 1 Screening of the decapeptide PS-SPCL for the ability to inhibit the binding of mAb 17/9 to the antigenic peptide Ac-YPYDVPDYASLRS-NH_2

Each panel represents one of the ten positional SPCLs (i.e. position-1 SPCL is Ac-O_1XXXXXXXXX-NH_2). Amino acids in the defined positions are represented on the *x*-axis of each panel by their single-letter code. The IC_{50} of the most effective peptide mixture at each position of the decapeptide library was normalized to 1.0 (*y*-axis), and the remaining peptide mixtures adjusted using the same factor.

standards were prepared by the divide, couple and recombine method (Houghten et al., 1991, 1992) and used as equimolar standards. Amino acid analysis showed the presence of each amino acid in approximately equimolar concentration (±25%). The cleavage and extraction of individual peptides and peptide mixtures from the resin was carried out as described (Houghten, 1985; Houghten et al., 1986, 1992). Purity and identity of individual peptides were characterized by reverse-phase h.p.l.c. and laser-desorption mass spectrometry, respectively. Assuming an average M_r of 1200 and a peptide mixture concentration of 10 mg/ml (8 mM), the concentration of each individual peptide within each peptide mixture is 40 fM.

E.l.i.s.a.

The competitive e.l.i.s.a. was carried out as described previously (Pinilla et al., 1993b). Briefly, 50 μl (2 μM) of the antigenic peptide Ac-YPYDVPDYASLRS-NH_2 was adsorbed to microtitre plates at 100 pmol/well. After blocking the plates with 1% BSA/PBS, 25 μl (8 mM) of each peptide mixture of the decapeptide PS-SPCL was added and serially diluted before adding 25 μl of mAb 17/9 to each well at a previously determined dilution. The concentration of peptide mixture necessary to inhibit 50% antibody binding (IC_{50}) was then determined. IC_{50} values were calculated using Graphpad (ISI Software, San Diego, CA, U.S.A.) and represent the average of at least three different determinations. Intra- and inter-assay variations were 10 and 25% respectively. Subsequent peptide mixtures and individual peptides were assayed in a similar manner.

RESULTS

Each of the ten positional peptide libraries making up this decapeptide PS-SPCL is composed of 20 peptide mixtures, in which a single position is defined with one of the 20 natural amino acids (represented as O), and the remaining nine positions of the 10-residue sequence are composed of mixtures (represented as X) of 18 amino acids (cysteine and tryptophan omitted). The ten positional peptide libraries have N-terminal acetyl and C-terminal amide groups and differ only in the location of their defined position. Each peptide mixture is made up of approx. 200 billion (2×10^{11}) individual sequences. Thus each positional set of 20 peptide mixtures, as well as the entire peptide library, is composed of approx. 4 trillion (4×10^{12}) decapeptides. This PS-SPCL can be represented as:

Position 1: Ac-O_1XXXXXXXXX-NH_2

Position 2: Ac-XO_2XXXXXXXX-NH_2

Position 3: Ac-XXO_3XXXXXXX-NH_2

Position 4: Ac-$XXXO_4$XXXXXX-NH_2

Position 5: Ac-$XXXXO_5$XXXXX-NH_2

Position 6: Ac-$XXXXXO_6$XXXX-NH_2

Position 7: Ac-$XXXXXXO_7$XXX-NH_2

Position 8: Ac-$XXXXXXXO_8$XX-NH_2

Position 9: Ac-$XXXXXXXXO_9$X-NH_2

Position 10: Ac-$XXXXXXXXXO_{10}$-NH_2

Each of the 200 peptide mixtures making up the decapeptide PS-SPCL was assayed by competitive e.l.i.s.a. for inhibition of the binding of mAb 17/9 to the antigenic peptide Ac-YPYDVPDYASLRS-NH_2. A graphical representation of the screening results is profiled in Figure 1. In order to compare the results of one position to another, the IC_{50} of the most effective peptide mixture for each position of the decapeptide library was normalized to 1.0, and the remaining peptide mixtures within that position were adjusted using the same factor. The IC_{50}

Table 1 IC_{50} values (mM) for peptide mixtures of each position of decapeptide PS-SPCL

Position 1		Position 2		Position 3		Position 4		Position 5	
D	0.10	D	0.05	D	0.03	D	0.07	D	0.02
E	0.35	E	0.13	E	0.22	E	0.19	Y	0.37
T	0.81	L	0.72	V	0.33	V	0.22	P	0.53
S	0.98	V	0.75	Y	0.51	P	0.34	V	0.55
Y	1.00	A	1.25	Q	1.06	I	0.56	E	0.56
G	1.04	T	1.41	P	1.23	Q	0.72	Q	0.96
L	1.04	P	1.51	I	1.37	K	0.84	I	1.20
V	1.07	S	1.55	L	1.45	T	1.32	K	1.29
H	1.15	Q	1.57	T	1.57	L	1.34	R	1.67
I	1.17	Y	1.58	R	1.63	Y	1.40	A	3.03
Q	1.48	W	1.66	W	1.67	R	1.43	T	4.00
N	1.71	I	1.70	K	1.71	H	1.48	G	4.00
K	1.74	C	1.94	A	2.56	A	1.62	W	4.00
A	1.81	N	2.80	C	2.84	W	2.50	F	4.00
P	1.88	M	2.92	M	3.10	C	2.68	M	4.00
W	1.88	R	3.10	H	4.00	F	2.68	S	4.00
M	1.98	K	4.00	S	4.00	M	2.72	N	4.00
F	4.00	H	4.00	N	4.00	N	2.91	H	4.00
C	4.00	G	4.00	F	4.00	S	2.93	C	4.00
R	4.00	F	4.00	G	4.00	G	4.00	L	4.00

Position 6		Position 7		Position 8		Position 9		Position 10	
D	0.06	E	0.08	D	0.02	Y	0.02	A	0.02
V	0.20	Y	0.09	A	0.08	A	0.17	S	0.06
I	0.23	P	0.10	S	0.26	D	0.24	R	0.53
Y	0.31	A	0.10	Y	0.31	S	0.65	Q	0.54
A	0.36	D	0.18	C	2.99	H	1.31	H	0.56
L	0.78	Q	0.73	G	3.09	G	1.92	D	0.66
E	1.18	K	0.96	H	3.14	E	2.35	G	0.76
S	1.25	S	0.98	R	4.00	W	2.74	T	0.76
T	1.28	R	1.57	N	4.00	C	2.86	N	0.77
K	1.43	V	2.60	W	4.00	V	4.00	E	0.83
Q	2.44	G	2.73	V	4.00	T	4.00	K	0.97
P	2.54	M	2.78	T	4.00	R	4.00	Y	1.35
M	2.57	C	2.90	E	4.00	M	4.00	P	1.39
C	2.65	N	4.00	F	4.00	Q	4.00	W	1.47
R	2.71	W	4.00	Q	4.00	P	4.00	L	1.49
H	2.73	F	4.00	P	4.00	N	4.00	I	1.52
W	2.78	T	4.00	M	4.00	K	4.00	V	1.57
F	2.86	I	4.00	K	4.00	I	4.00	M	2.09
N	4.00	H	4.00	I	4.00	F	4.00	F	2.09
G	4.00	L	4.00	L	4.00	L	4.00	C	4.00

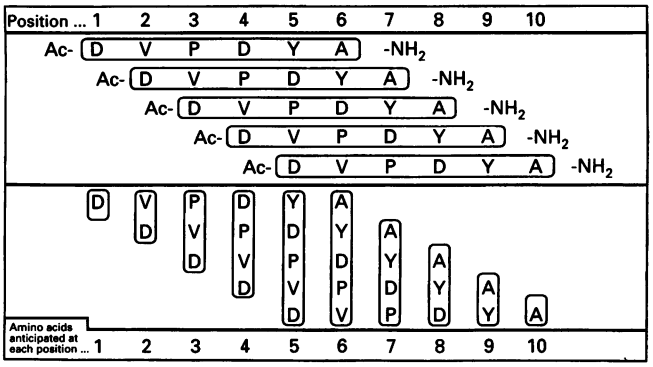

Figure 2 Expected overlapping series for six-residue antigenic determinant in decapeptide PS-SPCL and amino acids anticipated at each position

Table 2 Activities of -DVPDYA- 'walked through' ten positions of a mixture sequence

Peptide mixture	IC_{50} (nM)
Ac-XXXXDVPDYA-NH_2	1.6
Ac-XXXDVPDYAX-NH_2	6.3
Ac-XXDVPDYAXX-NH_2	5.2
Ac-XDVPDYAXXX-NH_2	5.5
Ac-DVPDYAXXXX-NH_2	8.2
XXXXDVPDYA-NH_2	1.6
XXXDVPDYAX-NH_2	7.5
XXDVPDYAXX-NH_2	6.8
XDVPDYAXXX-NH_2	12.1
DVPDYAXXXX-NH_2	78

values for each peptide mixture are shown in Table 1. For those peptide mixtures which did not inhibit more than 50% of the antibody binding to the 13-residue antigenic peptide, the highest concentration tested (4 mM) is shown. In the first and second position SPCL, only aspartic acid, and to a lesser extent glutamic acid, were found to effectively inhibit antibody binding, whereas the majority of the remaining peptide mixtures were found to have IC_{50} values of $\geqslant 1$ mM. For the third position SPCL, the aspartic acid peptide mixture was nearly seven times more effective at inhibiting antibody binding than the glutamic acid peptide mixture and ten times more effective than the next best peptide mixture (valine). The profile for the fourth position SPCL was less specific than the third position profile, with aspartic acid again the most effective peptide mixture, followed by glutamic acid, valine and proline. The fifth position SPCL was highly specific in that the aspartic acid peptide mixture was 18 times more effective than the tyrosine peptide mixture and nearly 25 times more effective than proline or valine; half of the peptide mixtures with the fifth position defined did not inhibit 50% antibody binding at the highest concentration tested (4 mM). Aspartic acid at the sixth position SPCL was the most effective peptide mixture: at least three times better than valine, isoleucine, alanine and tyrosine. In the seventh position, alanine, aspartic acid, glutamic acid, proline and tyrosine were found to have similar inhibitory activity. In the eighth position, only aspartic acid and, to a lesser extent, alanine, serine and tyrosine, was found to inhibit the binding of mAb 17/9. In the ninth position, tyrosine was eight and 12 times more effective than alanine and aspartic acid for antibody inhibition respectively. For the eighth and ninth position, there were 13 and 11 peptide mixtures respectively that did not inhibit 50% antibody binding at the highest concentration tested, which illustrates the high degree of antibody specificity for these two positions. Alanine and serine were approx. 10 times more effective than the other amino acids defined at the tenth position.

The results for the last six positions of the scanning profile of the decapeptide PS-SPCL when screened against mAb 17/9 are similar to the scanning profile obtained in an earlier study on screening a hexapeptide PS-SPCL composed of more than 34 million peptides. In the earlier study, the antigenic determinant -DVPDYA- recognized by mAb 17/9 was correctly identified directly from the data obtained in the initial library screening (Pinilla et al., 1992). Figure 2 illustrates the amino acids that would be anticipated at each position upon walking the same six-residue sequence (-DVPDYA-) through the 10 positions of the decapeptide PS-SPCL. The most effective peptide mixtures at each position in Table 1 are in agreement with the anticipated amino acids from Figure 2, with the most specific profile seen in positions 5–10. For example, the most effective amino acids at position 5 in Table 1 are the same as the amino acids that are anticipated in Figure 2, namely, aspartic acid, tyrosine, proline and valine. However, if there were no preference in the location of this antigenic determinant in a decapeptide for antibody recognition, then the activities of the peptide mixtures defined at positions 5–9 with tyrosine, which is known to be the most

Table 3 IC_{50} values for the 180 defined decapeptides derived from all combinations of amino acids found to be effective from the PS-SPCL screening

Position... 1 2 3 4 5 6 7 8 9 10
Amino acid... D D D D D A A A A A
V V D D D
E Y Y Y
P
Y

#		IC_{50} nM
1	Ac-D D D D D D A A A A -NH_2	8160
2	Ac-D D D D D D A A D A -NH_2	435000
3	Ac-D D D D D D A A Y A -NH_2	435000
4	Ac-D D D D D D A D A A -NH_2	435000
5	Ac-D D D D D D A D D A -NH_2	435000
6	Ac-D D D D D D A D Y A -NH_2	38
7	Ac-D D D D D D A Y A A -NH_2	435000
8	Ac-D D D D D D A Y D A -NH_2	435000
9	Ac-D D D D D D A Y Y A -NH_2	435000
10	Ac-D D D D D D D A A A -NH_2	435000
11	Ac-D D D D D D D A D A -NH_2	23400
12	Ac-D D D D D D D A Y A -NH_2	435000
13	Ac-D D D D D D D D A A -NH_2	435000
14	Ac-D D D D D D D D D A -NH_2	435000
15	Ac-D D D D D D D D Y A -NH_2	184
16	Ac-D D D D D D D Y A A -NH_2	203
17	Ac-D D D D D D D Y D A -NH_2	435000
18	Ac-D D D D D D D Y Y A -NH_2	435000
19	Ac-D D D D D D E A A A -NH_2	435000
20	Ac-D D D D D D E A D A -NH_2	435000
21	Ac-D D D D D D E A Y A -NH_2	435000
22	Ac-D D D D D D E D A A -NH_2	79000
23	Ac-D D D D D D E D D A -NH_2	380000
24	Ac-D D D D D D E D Y A -NH_2	8.7
25	Ac-D D D D D D E Y A A -NH_2	1873
26	Ac-D D D D D D E Y D A -NH_2	1732
27	Ac-D D D D D D E Y Y A -NH_2	176700
28	Ac-D D D D D D P A A A -NH_2	403300
29	Ac-D D D D D D P A D A -NH_2	435000
30	Ac-D D D D D D P A Y A -NH_2	435000
31	Ac-D D D D D D P D A A -NH_2	123900
32	Ac-D D D D D D P D D A -NH_2	28800
33	Ac-D D D D D D P D Y A -NH_2	8.0
34	Ac-D D D D D D P Y A A -NH_2	263600
35	Ac-D D D D D D P Y D A -NH_2	2949
36	Ac-D D D D D D P Y Y A -NH_2	202800
37	Ac-D D D D D D Y A A A -NH_2	414
38	Ac-D D D D D D Y A D A -NH_2	1890
39	Ac-D D D D D D Y A Y A -NH_2	3261
40	Ac-D D D D D D Y D A A -NH_2	54400
41	Ac-D D D D D D Y D D A -NH_2	22000
42	Ac-D D D D D D Y D Y A -NH_2	300
43	Ac-D D D D D D Y Y A A -NH_2	339800
44	Ac-D D D D D D Y Y D A -NH_2	351000
45	Ac-D D D D D D Y Y Y A -NH_2	33700
46	Ac-D D D V D D A A A A -NH_2	19100
47	Ac-D D D V D D A A D A -NH_2	16800
48	Ac-D D D V D D A A Y A -NH_2	67500
49	Ac-D D D V D D A D A A -NH_2	28100
50	Ac-D D D V D D A D D A -NH_2	2081
51	Ac-D D D V D D A D Y A -NH_2	62
52	Ac-D D D V D D A Y A A -NH_2	18000
53	Ac-D D D V D D A Y D A -NH_2	28700
54	Ac-D D D V D D A Y Y A -NH_2	30000
55	Ac-D D D V D D D A A A -NH_2	34800
56	Ac-D D D V D D D A D A -NH_2	53700
57	Ac-D D D V D D D A Y A -NH_2	32300
58	Ac-D D D V D D D D A A -NH_2	122800
59	Ac-D D D V D D D D D A -NH_2	106200
60	Ac-D D D V D D D D Y A -NH_2	281
61	Ac-D D D V D D D Y A A -NH_2	2774
62	Ac-D D D V D D D Y D A -NH_2	24300
63	Ac-D D D V D D D Y Y A -NH_2	74100
64	Ac-D D D V D D E A A A -NH_2	11100
65	Ac-D D D V D D E A D A -NH_2	45400
66	Ac-D D D V D D E A Y A -NH_2	2578
67	Ac-D D D V D D E D A A -NH_2	2849
68	Ac-D D D V D D E D D A -NH_2	3069
69	Ac-D D D V D D E D Y A -NH_2	6.7
70	Ac-D D D V D D E Y A A -NH_2	23900
71	Ac-D D D V D D E Y D A -NH_2	89600
72	Ac-D D D V D D E Y Y A -NH_2	51400
73	Ac-D D D V D D P A A A -NH_2	45800
74	Ac-D D D V D D P A D A -NH_2	81100
75	Ac-D D D V D D P A Y A -NH_2	69100
76	Ac-D D D V D D P D A A -NH_2	11200
77	Ac-D D D V D D P D D A -NH_2	266600
78	Ac-D D D V D D P D Y A -NH_2	4.4
79	Ac-D D D V D D P Y A A -NH_2	27400
80	Ac-D D D V D D P Y D A -NH_2	44200
81	Ac-D D D V D D P Y Y A -NH_2	10300
82	Ac-D D D V D D Y A A A -NH_2	6.3
83	Ac-D D D V D D Y A D A -NH_2	45
84	Ac-D D D V D D Y A Y A -NH_2	85
85	Ac-D D D V D D Y D A A -NH_2	14400
86	Ac-D D D V D D Y D D A -NH_2	6473
87	Ac-D D D V D D Y D Y A -NH_2	437
88	Ac-D D D V D D Y Y A A -NH_2	2757
89	Ac-D D D V D D Y Y D A -NH_2	50700
90	Ac-D D D V D D Y Y Y A -NH_2	37800
91	Ac-D D D D D V A A A A -NH_2	435000
92	Ac-D D D D D V A A D A -NH_2	227300
93	Ac-D D D D D V A A Y A -NH_2	109800
94	Ac-D D D D D V A D A A -NH_2	9500
95	Ac-D D D D D V A D D A -NH_2	85900
96	Ac-D D D D D V A D Y A -NH_2	0.8
97	Ac-D D D D D V A Y A A -NH_2	97700
98	Ac-D D D D D V A Y D A -NH_2	46200
99	Ac-D D D D D V A Y Y A -NH_2	57900
100	Ac-D D D D D V D A A A -NH_2	234800
101	Ac-D D D D D V D A D A -NH_2	276200
102	Ac-D D D D D V D A Y A -NH_2	108900
103	Ac-D D D D D V D D A A -NH_2	5300
104	Ac-D D D D D V D D D A -NH_2	9400
105	Ac-D D D D D V D D Y A -NH_2	2.8
106	Ac-D D D D D V D Y A A -NH_2	77
107	Ac-D D D D D V D Y D A -NH_2	45400
108	Ac-D D D D D V D Y Y A -NH_2	41300
109	Ac-D D D D D V E A A A -NH_2	57200
110	Ac-D D D D D V E A D A -NH_2	380500
111	Ac-D D D D D V E A Y A -NH_2	136300
112	Ac-D D D D D V E D A A -NH_2	3366
113	Ac-D D D D D V E D D A -NH_2	23000
114	Ac-D D D D D V E D Y A -NH_2	1.1
115	Ac-D D D D D V E Y A A -NH_2	626
116	Ac-D D D D D V E Y D A -NH_2	52700
117	Ac-D D D D D V E Y Y A -NH_2	29000
118	Ac-D D D D D V P A A A -NH_2	232100
119	Ac-D D D D D V P A D A -NH_2	435000
120	Ac-D D D D D V P A Y A -NH_2	78300
121	Ac-D D D D D V P D A A -NH_2	2405
122	Ac-D D D D D V P D D A -NH_2	13000
123	Ac-D D D D D V P D Y A -NH_2	0.6
124	Ac-D D D D D V P Y A A -NH_2	487
125	Ac-D D D D D V P Y D A -NH_2	49300
126	Ac-D D D D D V P Y Y A -NH_2	100600
127	Ac-D D D D D V Y A A A -NH_2	26300
128	Ac-D D D D D V Y A D A -NH_2	180600
129	Ac-D D D D D V Y A Y A -NH_2	112500
130	Ac-D D D D D V Y D A A -NH_2	38000
131	Ac-D D D D D V Y D D A -NH_2	10700
132	Ac-D D D D D V Y D Y A -NH_2	7.8
133	Ac-D D D D D V Y Y A A -NH_2	1832
134	Ac-D D D D D V Y Y D A -NH_2	55300
135	Ac-D D D D D V Y Y Y A -NH_2	34400
136	Ac-D D D V D V A A A A -NH_2	13300
137	Ac-D D D V D V A A D A -NH_2	63100
138	Ac-D D D V D V A A Y A -NH_2	11000
139	Ac-D D D V D V A D A A -NH_2	11300
140	Ac-D D D V D V A D D A -NH_2	45900
141	Ac-D D D V D V A D Y A -NH_2	2.2
142	Ac-D D D V D V A Y A A -NH_2	1290
143	Ac-D D D V D V A Y D A -NH_2	7200
144	Ac-D D D V D V A Y Y A -NH_2	14000
145	Ac-D D D V D V D A A A -NH_2	4900
146	Ac-D D D V D V D A D A -NH_2	38700
147	Ac-D D D V D V D A Y A -NH_2	14300
148	Ac-D D D V D V D D A A -NH_2	12500
149	Ac-D D D V D V D D D A -NH_2	7500
150	Ac-D D D V D V D D Y A -NH_2	9.5
151	Ac-D D D V D V D Y A A -NH_2	83
152	Ac-D D D V D V D Y D A -NH_2	42500
153	Ac-D D D V D V D Y Y A -NH_2	29700
154	Ac-D D D V D V E A A A -NH_2	81600
155	Ac-D D D V D V E A D A -NH_2	78900
156	Ac-D D D V D V E A Y A -NH_2	26100
157	Ac-D D D V D V E D A A -NH_2	3450
158	Ac-D D D V D V E D D A -NH_2	675
159	Ac-D D D V D V E D Y A -NH_2	3.3
160	Ac-D D D V D V E Y A A -NH_2	2253
161	Ac-D D D V D V E Y D A -NH_2	38600
162	Ac-D D D V D V E Y Y A -NH_2	30300
163	Ac-D D D V D V P A A A -NH_2	28600
164	Ac-D D D V D V P A D A -NH_2	141700
165	Ac-D D D V D V P A Y A -NH_2	17100
166	Ac-D D D V D V P D A A -NH_2	2725
167	Ac-D D D V D V P D D A -NH_2	18000
168	Ac-D D D V D V P D Y A -NH_2	1.1
169	Ac-D D D V D V P Y A A -NH_2	22700
170	Ac-D D D V D V P Y D A -NH_2	11500
171	Ac-D D D V D V P Y Y A -NH_2	31400
172	Ac-D D D V D V Y A A A -NH_2	6500
173	Ac-D D D V D V Y A D A -NH_2	25000
174	Ac-D D D V D V Y A Y A -NH_2	30700
175	Ac-D D D V D V Y D A A -NH_2	28900
176	Ac-D D D V D V Y D D A -NH_2	65800
177	Ac-D D D V D V Y D Y A -NH_2	7.6
178	Ac-D D D V D V Y Y A A -NH_2	13100
179	Ac-D D D V D V Y Y D A -NH_2	27800
180	Ac-D D D V D V Y Y Y A -NH_2	52800

specific residue of this antigenic determinant, should be the same. This was not the case, however, in that the peptide mixture having a defined tyrosine at position 9 was clearly more effective than any other position defined with tyrosine. Since this appeared to indicate that mAb 17/9 was preferentially recognizing -DVPDYA- at the C-terminal region of the decapeptide sequence, two new decapeptide mixture sets (N-acetylated and non-acetylated) were synthesized. In each set the sequence -DVPDYA- was 'walked through' the ten positions of a mixture sequence. Each of these peptide mixtures contains approx. 104976 peptides and was assayed by competitive e.l.i.s.a. The most effective peptide mixture was found to be Ac-XXXXDVPDYA-NH_2 (Table 2). This peptide mixture inhibited the binding of mAb 17/9 5 times more effectively than Ac-DVPDYAXXXX-NH_2. Similarly, the non-acetylated peptide mixture XXXXDVPDYA-NH_2 was found to have inhibiting activity equal to its acetylated form, but nearly 50 times more activity than the peptide mixture DVPDYAXXXX-NH_2. These results agree with the results found for the decapeptide PS-SPCL (Table 1 and Figure 1).

Individual peptides were synthesized in order to prove that the decapeptide PS-SPCL screening results permit the identification of highly active individual decapeptides. In the ideal case, the set of peptides synthesized would represent all of the possible sequence combinations from the most effective amino acid residues for each of the ten positions. However, the number of peptides that result from choosing more than two amino acids at each position rises very rapidly. For example, two amino acids for each position of the library would require the synthesis of $2^{10} = 1024$ individual peptides. To hold the number of peptide combinations to a minimum, two selection criteria were employed. Amino acids corresponding to the most effective peptide mixtures for each position were chosen on the basis of having chemically differing side chains, or IC_{50} values threefold or greater than the next best peptide mixture. Following these criteria, alanine was chosen over serine at the eighth, ninth and tenth positions, and aspartic acid was chosen over glutamic acid in positions 1–4. Valine was chosen for positions 4 and 6, since it is chemically different from aspartic acid. Since alanine, aspartic acid, glutamic acid, proline and tyrosine in the seventh position resulted in peptide mixtures with nearly equal activities, all five were included. Aspartic acid and alanine were chosen for position eight, since they are chemically distinct. Although the tyrosine peptide mixture at position eight was found to be nearly 15-fold less effective than the aspartic acid peptide mixture, tyrosine was also included for this position, since its activity was 10-fold greater than the remaining residues defined at this position. For the same reason, aspartic acid was included with tyrosine and alanine at position 9. In summary, the amino acid residues chosen were the following: aspartic acid for positions 1–3; aspartic acid and valine for position 4; aspartic acid for position 5; aspartic acid and valine for position 6; alanine, aspartic acid, glutamic acid, proline and tyrosine for position 7; alanine, aspartic acid and tyrosine for positions 8 and 9; and alanine for position 10. All of the possible combinations of these amino acids ($1 \times 1 \times 1 \times 2 \times 1 \times 2 \times 5 \times 3 \times 3 \times 1$) yield a total of 180 individual N-acetylated decapeptides.

This set of individual peptides was synthesized and assayed against mAb 17/9 by competitive e.l.i.s.a. (results are shown in Table 3). To facilitate interpretation of the data, this Table is arranged in four columns of sequences, namely nos. 1–45, 46–90, 91–135 and 136–180. When the peptides are compared horizontally, the only difference between these groups occurs at positions 4 and 6. Peptides 1–45 have aspartic acid at these two positions. Peptides 46–90 differ from this first group by a single valine substitution at position 4 (shaded). For example, peptides 1 and 46 have the same sequence, except for position four, in which the aspartic acid is substituted by valine. Peptides 91–135 differ from the first group by a single valine substitution at position 6 (shaded). Finally, peptides 136–180 differ from the first group by valine substitutions at positions 4 and 6 (shaded). In this manner, the activities of single or double substitution analogues of the peptides 1–45 can be readily seen; 15 peptides were found to have IC_{50} values comparable with the antigenic peptide Ac-YPYDVPDYASLRS-NH_2 ($IC_{50} = 6$ nM) and are outlined in Table 3. All but one of these highly active peptides (namely 82) have -DYA- at positions 8–10, which correspond to the C-terminus of the known six-residue antigenic determinant. Four of these peptides (96, 114, 123 and 168), all of which contain multiple aspartic acids in the N-terminal region, were found to be nearly five times more effective than the 13-residue control peptide. Also, it is seen that the most effective peptides having valine at position 6 have nearly a 10-fold better activity than those with aspartic acid. Of the 180 decapeptides assayed, 150 were found to have IC_{50} values greater than 1 μM. These peptides were therefore at least 200 times less effective than the antigenic peptide Ac-YPYDVPDYASLRS-NH_2 ($IC_{50} = 6$ nM). Of these 150 relatively poorly active peptides, only three, namely peptide 38 ($IC_{50} = 1890$ nM), 61 ($IC_{50} = 2774$ nM) and 39 ($IC_{50} = 3281$ nM) were found to contain the three-residue motif -DYA-. This motif is known to play a key role in this antigen–antibody interaction (Rini et al., 1992). Conversely, 21 decapeptides having IC_{50} values lower than 100 nM contain this motif. More importantly, 17 peptides have the -DYA- motif in positions 8–10. Finally, it should be noted that if only the best peptide mixture from each position in Table 1 were selected, the corresponding decapeptide would result in a highly active sequence, namely peptide 24 in Table 3. Ac-DDDDDDEDYA-NH_2 was found to have an IC_{50} of 8.7 nM; the aspartic acid and glutamic acid residues in the sixth and seventh positions respectively represent acceptable substitutions for the redundant valine and proline residues of the antigenic determinant -DVPDYA-.

DISCUSSION

The decapeptide PS-SPCL described here was designed as a tool to study molecular interactions having recognition sites made up of more than six residues. In order to evaluate the validity and practical utility of a peptide library made up of approx. 4 trillion (4×10^{12}) peptides, this decapeptide PS-SPCL was screened by competitive e.l.i.s.a. against a well-characterized peptide–antibody interaction. The fine specificity of this peptide–mAb interaction has been examined using a complete series of single substitution analogues (19 peptide analogues for each position) of the antigenic determinant of the antigenic peptide Ac-YPYDVPDYASLRS-NH_2 (Pinilla et al., 1993b). This interaction has also been determined by X-ray crystallography (Rini et al., 1992), and the results are in good agreement with the e.l.i.s.a. substitution profile found for the antigenic determinant. The screening of the decapeptide PS-SPCL against this peptide–antibody interaction was also of value in determining the influence of the flanking regions of the antigenic determinant on antibody recognition. The results indicate that the six-residue antigenic determinant is preferentially recognized by mAb 17/9 at the C-terminal region of the decapeptide.

The utility of the PS-SPCL approach is that each position of the peptide library can be screened simultaneously and, if the results can be clearly differentiated from one another and the position to position connectivity established, then the iterative synthesis and selection steps that are necessary for the dual-fixed

position SPCLs (Houghten et al., 1991, 1992) are not required for the identification of highly active peptides. Individual peptides derived from the combinations of the amino acids of the best peptide mixtures from each position can then be synthesized to confirm the library screening results. In the current study, the results of the library screening prompted the synthesis of a set of 180 individual decapeptides. The most effective peptides were found to contain the six-residue antigenic determinant (or conservative analogue of the same) preceded by four aspartic acid residues; four decapeptides were found to have higher affinities relative to the known antigenic determinant. These results show that one can identify peptides with high affinity from a peptide library of nearly 4 trillion (4×10^{12}) sequences. It should be noted that no information about the sequence of the antigen or antibody is required to carry out determinations of this kind.

As an alternative to the immediate synthesis of individual peptides following the screening of a PS-SPCL, as was carried out here, one can prepare a sub-library in the same scanning format composed of only those amino acids found to be effective or which fall below a suitable IC_{50} cut-off. This may be a preferred approach for those interactions screened in which no sequence information is available, or for those cases in which the screening profile is not clear. In this manner, the number of individual peptides in each peptide mixture is substantially reduced. For example, if only five amino acids are used at each position for a decapeptide, then the number of peptides in each peptide mixture will decrease from 200 billion (2×10^{11}) to approx. 2 million. This would in turn increase the effective concentration of each individual peptide within each peptide mixture, and may lead to more specific results. Upon screening of this sub-library, it may then be easier to select the most effective amino acid at each position, which minimizes the number of individual peptides needed to identify the most effective decapeptide. This approach would require more time due to the additional synthesis, but can be expected to reduce the number of individual peptides ultimately synthesized.

The results from screening the decapeptide library clearly show that antibody binding is more favourable with the antigenic determinant located at the C-terminal than at the N-terminal. From the structural information of this peptide–antibody interaction, it is known that four residues of the antigenic determinant (-DYAS-) adopts a β-turn when bound to the antibody (Rini et al., 1992). Thus extension of the C-terminal region of the antigenic determinant with 1–4 mixture positions may result in disruption of the β-turn conformation, resulting in a decrease in the overall binding energy of this peptide–antibody interaction. As anticipated from the screening results, the most effective peptides corresponded to the antigenic determinant at the C-terminal region with multiple aspartic acid residues in the N-terminal region. These aspartic acid residues, and possibly to a lesser extent the glutamic acid residues, in positions 1–4 can contribute to the binding of the antigenic determinant sequence to the antibody through weak ionic interactions with one or more antibody residues outside the antibody-binding pocket.

It should be noted that the number of individual peptides making up each peptide mixture in the decapeptide PS-SPCL is much greater than the number of peptides in each peptide mixture of either the hexapeptide PS-SPCL or the dual fixed-position SPCL (approx. 200 million versus 1.8 million and 130321 respectively). Thus, assuming an average M_r of 785 for each of the 400 peptide mixtures of Ac-O_1O_2XXXX-NH_2, at a concentration of 10 mg/ml, each individual peptide is present at 100 nM. Assuming the same M_r and concentration for each peptide mixture of the hexapeptide PS-SPCL, each peptide is present at a concentration of 7 nM. Assuming an average M_r of 1200 for the decapeptide PS-SPCL, and a peptide mixture concentration of 10 mg/ml, the concentration of each individual peptide is 40 fM. However, if the recognition site of an interaction is made up of only six residues, as is the interaction studied here, then the binding threshold or 'effective' concentration of the hexamer sequence within a decapeptide, is approx. 100000 times higher. The binding threshold of -DVPDYA- when present within a decapeptide is therefore approx. 4 nM, which is well within the detection range for this peptide–antibody interaction. To simplify this concept, if one envisions that the interaction of the antigenic determinant side chains of only the aspartic acids at positions 1 and 4, the tyrosine at position 5, and the alanine at position 6 with the residues of the antibody-binding site are responsible for the binding energy of this interaction, then six out of the ten positions in a decapeptide are non-specific. This equates to an effective concentration of the four-residue -D--DYA- motif within a decapeptide of not 40 fM, but 18^6 higher, or approx. 1.4 μM. It should also be noted that, at some point in an ever-increasing diversity pool, one will not have even a single molecule of each peptide at a workable concentration of mixtures. Thus, although 200 billion (2×10^{11}) peptides per mixture may seem too large a number for reasonable evaluation, a sufficient number of copies of each peptide are present. For example, at 10 mg/ml, a 50 μl sample contains approx. 1 million copies of each peptide making up each peptide mixture.

The general utility of this and other SPCLs is based on the fact that: (1) the peptides making up each mixture are present in approximately equimolar amounts; (2) the peptides are soluble (i.e., not support-bound), which allows each peptide to interact freely in solution with relevant receptors; and (3) peptide motifs are present at an effective concentration expected to yield a detectable signal in the majority of *in vitro* bioassays. It should be noted that, since the PS-SPCL described here is composed of ten separate positional SPCLs, each one can be considered independent of the others. Therefore, if desired, each positional SPCL can be independently screened and pursued using the iterative synthesis and selection process described earlier (Houghten et al., 1991, 1992). We have found that, when used in concert, the PS-SPCLs and dual fixed position SPCLs (i.e. Ac-O_1O_2XXXX-NH_2) have broad utility and application for the efficient identification of high-affinity bioactive peptides in basic research and drug discovery.

We thank James Winkle and Vince Hamashin for library synthesis, Edward Brehm for peptide analysis, Sandry Chendra and Jaime Buencamino for e.l.i.s.a. work, Dr. Ian Wilson for supplying monoclonal antibody 17/9, and Eileen Silva for her assistance in preparing this manuscript. This work was funded by Houghten Pharmaceuticals Inc., San Diego, CA, U.S.A.

REFERENCES

Appel, J. R., Pinilla, C. and Houghten, R. A. (1992) Immunomethods **1**, 17–23
Blondelle, S. E. and Houghten, R. A. (1994) in Techniques in Protein Chemistry V (Crabb, J. W., ed.), pp. 509–516, Academic Press, Orlando, in the press
Cull, M. G., Miller, J. F. and Schatz, P. J. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 1865–1869
Cwirla, S. E., Peters, E. A., Barrett, R. W. and Dower, W. J. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6378–6382
Devlin, J. J., Panganiban, L. C. and Devlin, P. E. (1990) Science **249**, 404–406
Dooley, C. T., Chung, N. N., Schiller, P. W. and Houghten, R. A. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 10811–10815
Dooley, C. T. and Houghten, R. A. (1993) Life Sci. **52**, 1509–1517
Eichler, J. and Houghten, R. A. (1993) Biochemistry **32**, 11035–11041
Geysen, H. M., Rodda, S. J. and Mason, T. J. (1986) Mol. Immunol. **23**, 709–715
Houghten, R. A. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5131–5135
Houghten, R. A. and Dooley, C. T. (1993) BioMed. Chem. Lett. **3**, 405–412

Houghten, R. A., Bray, M. K., De Graw, S. T. and Kirby, C. J. (1986) Int. J. Peptide Protein Res. **27**, 673–678
Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. and Cuervo, J. H. (1991) Nature (London) **354**, 84–86
Houghten, R. A., Appel, J. R., Blondelle, S. E., Cuervo, J. H., Dooley, C. T. and Pinilla, C. (1992) Biotechniques **13**, 412–421
Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. (1991) Nature (London) **354**, 82–84

Received 17 January 1994/24 February 1994; accepted 9 March 1994

Lowman, H. B., Bass, S. H., Simpson, N. and Wells, J. A. (1992) Biochemistry **30**, 10832–10838
Pinilla, C., Appel, J. R., Blanc, P. and Houghten, R. A. (1992) Biotechniques **13**, 901–905
Pinilla, C., Appel, J. R. and Houghten, R. A. (1993a) Gene **128**, 71–76
Pinilla, C., Appel, J. R. and Houghten, R. A. (1993b) Mol. Immunol. **30**, 577–585
Rini, J. M., Schulze-Gahmen, U. and Wilson, I. A. (1992) Science **255**, 959–965
Wilson, I. A., Niman, H. L., Houghten, R. A., Cherenson, A. R., Connolly, M. L. and Lerner, R. A. (1984) Cell **37**, 767–778