

Immunogenically Fit Subunit Vaccine Components Via Epitope Discovery from Natural Peptide Libraries¹

Leslie J. Matthews,² Robert Davis, and George P. Smith

Antigenic peptides that bind pathogen-specific Abs are a potential source of subunit vaccine components. To be effective the peptides must be immunogenically fit: when used as immunogens they must elicit Abs that cross-react with native intact pathogen. In this study, antigenic peptides obtained from phage display libraries through epitope discovery were systematically examined for immunogenic fitness. Peptides selected from random peptide libraries, in which the phage-displayed peptides are encoded by synthetic degenerate oligonucleotides, had marginal immunogenic fitness. In contrast, 50% of the peptides selected from a natural peptide library, in which phage display segments of actual pathogen polypeptides, proved very successful. Epitope discovery from natural peptide libraries is a promising route to subunit vaccines. *The Journal of Immunology*, 2002, 169: 837–846.

For many important infectious diseases, including malaria, conventional killed or attenuated vaccines are impractical. Subunit vaccines, consisting of pathogen-derived Ags, offer hope of effective, safe, and inexpensive protection. A few subunit vaccines consisting of whole polypeptides have proven to be successful—notably, recombinant hepatitis B surface Ag and tetanus toxoid. Even when no suitable whole polypeptide Ags are known, shorter fragments of pathogen proteins may suffice for subunit vaccines (1–6). In this paper, the term peptide is used regardless of the number of amino acids.

Candidate peptide vaccine components representing B cell epitopes are typically identified by their ability to bind Abs from subjects exposed to the relevant pathogen, which we call direct Abs. To be useful as a vaccine component, a peptide must be not only antigenic but also immunogenically fit: when used as an immunogen, the indirect Abs it elicits must cross-react with native intact pathogen. Immunogenic fitness is gauged by the fraction of indirect anti-peptide Abs that cross-react with the pathogen. It is distinct from immunogenicity, which is gauged by the total anti-peptide titer of those indirect Abs, including Abs that do not cross-react with pathogen. Although both immunogenicity and immunogenic fitness contribute to overall protection, the work reported in this paper focuses specifically on immunogenic fitness.

Peptides with excellent antigenicity and immunogenicity frequently lack adequate immunogenic fitness and therefore fail as potential vaccine components (7–15). A common explanation for this poor immunogenic fitness is the conformational flexibility of most short peptides. A flexible peptide may bind well to direct Ab and thus have good antigenicity; indeed, flexibility may sometimes enhance antigenicity by allowing the peptide to bind by an induced fit mechanism (16, 17). Likewise, a flexible peptide may be highly immunogenic, eliciting substantial Ab titers. However, if the peptide has a large repertoire of conformations, a preponderance of

those Abs may fail to recognize the corresponding native epitope on the intact pathogen (17–20).

Despite the importance of immunogenic fitness, it is rarely feasible to evaluate it in isolation from other vaccine qualities. In this paper we report a systematic investigation of immunogenic fitness using bacteriophage T4 as a model pathogen and mice as model patients. Although T4 does not cause disease in mice, it is a good surrogate for a complex pathogen, having nearly 30 surface-exposed proteins that are foreign to mammals. Its virtue as a model is that it can be safely and copiously prepared in pure form for use as an immunochemical reagent, allowing direct quantitation of immunogenic fitness. Because immunogenic fitness is a property of epitope structure and the response characteristics of immune cells, and does not depend on idiosyncratic details of pathogenesis, the results of this study should be applicable to pathogenic agents of medical importance.

Antigenic peptides were obtained through a strategy called epitope discovery (21), in which direct Ab is used to affinity select Ags from very large libraries of peptides displayed on filamentous phage carriers (22, 23). The property that enables a peptide to prevail during selection—high affinity for a prevalent subspecificity in the selecting direct Ab population—augurs well for success as a candidate peptide vaccine component, even though its correlation with immunogenic fitness is imperfect.

Selections were made from two types of libraries: random peptide libraries (RPLs),³ in which the phage-displayed peptides are encoded by synthetic random degenerate oligonucleotide inserts (24–28); and a natural peptide library (NPL), in which the phage particles display fragments of natural pathogen proteins, encoded by short DNA fragments of the pathogen genome (the T4 chromosome in our model system). Libraries of natural peptides representing single genes or antigenic regions have been used previously for mapping antigenic (29–32) and immunogenic (33) epitopes. However, to our knowledge, this is the first attempt to survey an entire genome for immunogenic peptides, using an NPL. Ligands affinity selected from RPLs and NPLs will be called random antigenic peptides (RAPs) and natural antigenic peptides (NAPs), respectively. We show that, while RAPs have only marginal immunogenic fitness, a large fraction of NAPs have excellent

Division of Biological Sciences, University of Missouri, Columbia, MO 65211

Received for publication June 14, 2001. Accepted for publication May 8, 2002.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ This work was supported by U.S. Army Grant DAAL03-92-G-0178 and National Institutes of Health Grant GM41478.

² Address correspondence and reprint requests to Dr. Leslie J. Matthews, Division of Biological Sciences, Tucker Hall, University of Missouri, Columbia, MO 65211-7400. E-mail address: matthewslj@missouri.edu

³ Abbreviations used in this paper: RPL, random peptide library; NPL, natural peptide library; RAP, random antigenic peptide; NAP, natural antigenic peptide; PEG, polyethylene glycol; MBP, maltose binding protein.

immunogenic fitness. We argue that epitope discovery using NPLs is a highly promising route to peptide vaccines.

Materials and Methods

Standard solutions

Standard solutions TE, BSA, dialyzed BSA, TBE, TBS, TBS/Tween (TBS/Tween supplemented with dialyzed BSA and azide), and *N*-Z-amine and yeast extract liquid and agar media were prepared as described (23), as was Dulbecco's PBS (D-PBS) (34).

Bacteria and phages

Escherichia coli K-12 strain K91BlueKan (23) is Hfr Cavalli with chromosomal genotype *lacZΔM15 lacY::mkh lacI^R thi*; the engineered *mkh* transposon confers resistance to kanamycin. K-12 strain MC1061 (35) (W. Dower, Affymax, Palo Alto, CA) is F⁻ with chromosomal genotype *hsdR mcrB Δ(araABC-leu)6779 araD139 Δlac174 galU galK strA thi*.

Filamentous phage clones were routinely propagated in strain K91BlueKan and cultured in NZY containing 20 μg/ml tetracycline. For clones derived from libraries constructed in the f88-4 vector (Table I), 1 mM isopropyl-β-D-thiogalactoside was included in the growth medium to fully induce expression of the fusion protein, which is transcribed from a *tac* promoter.

Filamentous phage were partially purified from culture supernatant by two polyethylene glycol (PEG) precipitations as described (23). PEG-precipitated virions were further purified as required by CsCl equilibrium density gradient centrifugation (22). Phage for mouse immunizations were purified by detergent extraction, PEG precipitation, and CsCl equilibrium density gradient centrifugation (36).

Wild-type T4D was obtained from F. Eiserling (University of California, Los Angeles, CA). T4D amber mutant T4*amE727J* (37) lacks the Wac protein and was obtained from W. Wood (University of Colorado, Boulder, CO). Amber mutant T4*Dhoc* (38), lacking the Hoc protein, and T4B mutant T4eG326 (39–41) were provided by L. Black (University of Maryland, Baltimore, MD). T4B mutant T4eG326, whose deletion spans genes *ipii* and *ipiii* (encoding internal proteins IPII and IPIII) and part of gene *e* (encoding lysozyme), served as the IPIII-less form of T4 in ELISAs. Except for T4eG326, general procedures for enumerating and propagating T4 from a single plaque were as described (42). The lysozyme-less mutant T4eG326 was enumerated on nutrient agar petri dishes supplemented with chicken egg-white lysozyme (43). Large batches of T4D wild-type, T4*amE727J*, and T4*Dhoc* virions were propagated as described (44); large batches of T4eG326 virions were propagated in a series of one-step growth experiments (43). T4 virions were purified by sucrose density gradient centrifugation.

NPL construction

A detailed description of the procedure for NPL construction can be found on our web site (<http://www.biosci.missouri.edu/SmithGP/index.html>). Unmodified T4 DNA (42) was digested with serial dilutions of DNase I (Boehringer Mannheim, Indianapolis, IN) in Mn²⁺ buffer (45). Digests whose peak fragment size (estimated by PAGE) was ≥100 bp were pooled and polished with T4 DNA polymerase (Boehringer Mannheim) followed by exonuclease⁻ DNA polymerase I Klenow fragment (Promega, Madison, WI) (46). Polished DNase I-digested DNA fragments were ligated to

5'-phosphorylated, blunt-end, hairpin linkers containing *Hind*III (5'-pAGCGGCAAAGCTTCGGTGCACGGAGAATACCTCCGTGCACCG AAGCTTTGCCGCT) and *Pst*I (5'-pCGCTGCAGGACCTGGTTCCGA ATACCGGAACCAGGTCTGCAGCG) restriction sites. The linker-ligated fragments were digested with exonuclease III (Invitrogen, Carlsbad, CA) and exonuclease VII (Invitrogen) to degrade fragments that were not successfully linker-ligated at both ends; cleaved with *Hind*III and *Pst*I (Promega); fractionated by PAGE to remove the short end fragments; and spliced into vector f88-4 (S. Choukri, unpublished observation; GenBank accession no. AF218363; also cut with *Hind*III and *Pst*I), which displays up to ~150 guest peptides fused to the major coat protein pVIII. After ligation, the DNA was ethanol precipitated and concentrated to a final volume of ~50 μl on a Centricon 30-kDa ultrafilter (Millipore, Bedford, MA). Aliquots were electroporated into MC1061 cells and amplified as described (23).

Direct and indirect antisera and Abs

Ten BALB/c mice (Charles River Breeding Laboratories, Wilmington, MA) were immunized with 100 μg purified T4D (2.7 × 10¹¹ particles in D-PBS) administered s.c. weekly for 6 wk. Mice were exsanguinated 1 wk after the final immunization. The resulting direct antisera were stored at -20°C. Direct IgG Ab was purified from pooled antisera by protein A/G affinity chromatography (Pierce, Rockford, IL) and biotinylated with Biotin-XX-NHS (Molecular Probes, Eugene, OR) according to the manufacturers' recommendations. The resulting biotinylated protein will be called direct Bio-IgG.

Indirect antiserum was prepared in BALB/c mice (Charles River Breeding Laboratories) as described (47). Preimmune serum was obtained immediately before the first injection. The mice were injected i.p. three times at 3-wk intervals with 10¹¹ peptide-bearing filamentous phage (detergent/CsCl purified) or 20 μg IPIII fusion protein in D-PBS and emulsified in an equal volume of IFA by 100–200 passages through an 18-gauge double-hub needle. Negative control mice were injected with the same carriers (wild-type fd phage or IPIII protein) bearing no peptide. Mice were exsanguinated 10 days after the final immunization and the resulting indirect antisera were stored at -20°C.

Direct Bio-IgG and indirect antiserum elicited by phage immunization were absorbed with wild-type fd phage particles to remove any traces of Abs that react with the phage carrier. Bio-IgG (160 μg) was mixed with 4 × 10¹³ fd virions (~1 mg phage protein) in D-PBS or TTDBA buffer. Serum (20 μl) was mixed with 8 × 10¹³ fd virions (~2 mg phage protein) in D-PBS to give an overall dilution of 1/40 relative to the original serum. After overnight incubation at 4°C, the mixtures were centrifuged in a Beckman TLA100.3 rotor (Beckman Coulter, Fullerton, CA) at 57,000 rev/min for 50 min at 4°C to pellet phage along with any bound Abs. The supernatants were transferred to fresh tubes, centrifuged again as described above, transferred to fresh tubes, and stored at 4°C. Preabsorption with wild-type fd phage was not necessary for indirect antisera elicited by IPIII-displayed peptides.

Immunoabsorption with T4

For assessing immunogenic fitness, samples of all indirect antisera were immunoabsorbed with T4. Two matched aliquots of indirect antisera were diluted in D-PBS: either 150 μl of fd-absorbed antiserum at a 1/40 dilution

Table I. Phage display libraries used in this study

Library (Ref.)	GenBank Accession No.	Type of Library	No. of Clones	Vector	Sequence of Displayed Peptide	Type of Constraint	Copy No.	Host Coat Protein
f3-6mer (27)	AF246446	RPL	2 × 10 ⁸	fUSE5	ADGAX ₆	None	5	pIII
f3-15mer (71)	AF246445	RPL	2.5 × 10 ⁸	fUSE5	ADGAX ₁₅	None	5	pIII
f88-15mer ^a	AF246448	RPL	2 × 10 ⁹	f88-4	X ₁₅	None	~150	pVIII
f88-Cys0 ^b	AF246449	RPL	5.6 × 10 ⁸	f88-4	AX ₅ CCX ₅	Disulfide	~150	pVIII
f88-Cys1 ^b	AF246450	RPL	2.8 × 10 ⁹	f88-4	AX ₅ CXCX ₅	Disulfide	~150	pVIII
f88-Cys2 ^b	AF246451	RPL	5.5 × 10 ⁷	f88-4	AX ₅ CX ₂ CX ₅	Disulfide	~150	pVIII
f88-Cys3 ^b	AF246452	RPL	5.2 × 10 ⁷	f88-4	AX ₅ CX ₃ CX ₄	Disulfide	~150	pVIII
f88-Cys4 ^b	AF246453	RPL	1.7 × 10 ⁸	f88-4	AX ₄ CX ₄ CX ₅	Disulfide	~150	pVIII
f88-Cys5 ^b	AF246454	RPL	5.9 × 10 ⁸	f88-4	AX ₄ CX ₅ CX ₄	Disulfide	~150	pVIII
f88-Cys6 ^b	AF246455	RPL	2.7 × 10 ⁸	f88-4	AX ₄ CX ₆ CX ₄	Disulfide	~150	pVIII
f8-8mer (72)	AF246447	RPL	2 × 10 ⁹	f8-1	AX ₈	Close packing	3,900	pVIII
T4		NPL	2.2 × 10 ⁸	f88-4	T4 peptides	None	~150	pVIII

^a S. Choukri, unpublished observation.

^b G. P. Smith, unpublished observation.

(~40 μg total IgG) or 50 μl of non-fd-absorbed antiserum at a 1/20 dilution (~25 μg total IgG). A suspension of 200 μl T4D in D-PBS ($4\text{--}6 \times 10^{12}$ particles/ml; 150–225 μg total T4 protein) was added to fd-absorbed antisera, or a suspension of 225 μl T4D in D-PBS (2.5×10^{12} particles/ml; 100 μg total T4 protein) was added to non-fd-absorbed antisera. For mock absorption, an equal volume of D-PBS or an equal amount and concentration of a mutant form of T4 missing the relevant protein was added. T4D_{hoc} was used to absorb antisera elicited by the Hoc 1–89 and Hoc 320–347 NAPs and T4amE727J was used to absorb antisera elicited by the Wac 461–487 NAP. After overnight incubation at 4°C, the T4-absorbed and mock-absorbed antisera were centrifuged at 13,000 rev/min for 30 min in a microcentrifuge. The supernatants were transferred to fresh microcentrifuge tubes, centrifuged again as before, transferred to fresh microcentrifuge tubes, and stored at 4°C. None of the T4-absorbed antisera showed residual reactivity against T4 in ELISAs (data not shown). The same procedure was used to prepare T4-absorbed and mock-absorbed direct Bio-IgG for assessing pathogen specificity.

Affinity selection

Direct anti-T4 Bio-IgG that had been preabsorbed with wild-type fd phage was used to affinity select phage-borne peptides from each of the 12 phage display libraries (Table I) by the one-step method (23). Yields were quantified and phage eluates were amplified as described (23). To avoid selecting streptavidin-binding phages, we alternated between immobilizing the Bio-IgG onto the plastic surface with streptavidin vs neutravidin (Pierce) in consecutive rounds of selection. In addition, streptavidin or neutravidin molecules not bound to Bio-IgG were blocked with biotin before adding phage libraries. No streptavidin-binding phage emerged from the selections.

Screening affinity selection outputs

RAPs (5–10 from each affinity selection final output; 120 total) and NAPs (68 total) were randomly chosen from the final affinity selection outputs, propagated on the small scale (23), partially purified by PEG precipitation, and screened by ELISAs in which the immobilized phage were reacted with direct Bio-IgG or antiserum. Based on the ELISA screening results, 43 phage clones (Table II) with relatively high Ab binding activity were chosen from the RPLs for further characterization as described (23), including at least two clones from each library except Cys2. Sixty-eight phage clones from the NPL that showed relatively high Ab binding activity by ELISA were screened by one-lane sequencing (23) to identify groups of clones with identical inserts. Clones representing 15 unique inserts were further characterized by complete sequencing, yielding the NAPs listed in Table III.

IPIII and MBP fusion proteins

A subset of affinity-selected peptides were fused to both maltose binding protein (MBP) and His-tagged IPIII fusion partners using the pET-29a⁺ vector (Novagen, Madison, WI). The IPIII fusion constructs included the following (in order): the 6-bp vector *NdeI* site (including the ATG start codon); the coding sequence for the square-bracketed amino acids in Tables II and III; the reverse complement of T4 nucleotides 65934–66382 (GenBank accession no. AF158101.3), encoding the entirety of the mature form of the IPIII protein; and the 6-bp vector *XhoI* site, which is followed by the six codons for the His tag. The MBP fusion constructs included the following (in order): the 6-bp vector *NdeI* site; the coding sequence for the square-bracketed amino acids in Tables II and III; the coding sequence (GCTTCTCTGGTGCCACGCGGC) for a thrombin cleavage site; the reverse complement of *E. coli* nucleotides 11939–13065 from GenBank accession no. AE000476, encoding the last four amino acids of the signal peptide and the entire mature form of MBP, and including the MBP stop codon; and the 6-bp vector *XhoI* site. Fusion proteins were expressed according to the supplier's instructions (Novagen) and were extracted in B-PER lysis solution (Pierce) according to the supplier's instructions. The His-tagged IPIII fusion proteins were affinity-purified on nickel affinity columns (6 \times His fusion protein purification kit; Pierce) in the presence of 7 M guanidinium chloride; MBP fusion proteins were affinity-purified on amylose columns (New England Biolabs, Beverly, MA) and biotinylated as previously described. Proteins were quantified spectrophotometrically in 6 M guanidinium chloride (48).

Anti-peptide titer and percentage of cross-reactivity

Wells of ELISA dishes were coated with 5×10^{10} of the corresponding filamentous phage particles (purified by either PEG precipitation or CsCl ultracentrifugation) in 50 μl TBS for 1–2 h at room temperature, washed with TBS/Tween, and reacted with 100–200 μl of T4-absorbed or mock-

Table II. Random antigenic peptides affinity selected with direct anti-T4 Ab

No.	Displayed Peptide ^a	No. of Clones ^b	Library
1	SEWSPPHRGTAFFPS	1	f88-15mer
2	QEWNPWRLLTLFVTR	1	f88-15mer
3	aL SEWCPPHRTc NNKS	1	f88-Cys5
4	a PMEWcPPHRPc TPPL	2	f88-Cys5
5	[a PTEWCPPHRTc WPTTpa ^e] ^c	1	f88-Cys5
6	a TQWCPPHRQc SPIP	1	f88-Cys5
7	a ESEWCPPTRPgc TEPW	1	f88-Cys6
8	adga RLGHFTFFPY APYA	1	f3-15mer
9	a GAEFFYAS	1	f8-8mer
10	a EFFYEGSK	1	f8-8mer
11	SADNGT EFFYPP ILLA	1	f88-15mer
12	TGSMSS EFFYAP SPS	1	f88-15mer
13	AEFFYSEY SVPVTPPL	1	f88-15mer
14	a YFFYDcc PAGTY	1	f88-Cys0
15	a EFFYERYPTFQ T	1	f88-Cys0? ^d
16	a EFFYc YEPDcSKWN	1	f88-Cys4
17	a EFFYc AETKcFEMA	1	f88-Cys4
18	a EFFYc LKDTcVLAP	1	f88-Cys4
19	a EFFYc TEKQcWSRP	1	f88-Cys4
20	a YFFYc ETPPcHYAP	1	f88-Cys4
21	a EFPLc SAEPNcTKAE	1	f88-Cys5
22	adga AFWWGY	5	f3-6mer
23	adga NFWWGY	3	f3-6mer
24	a EMNYG PVS	1	f8-8mer
25	a EMNYP DRS	1	f3-8mer
26	adga DFSIRA	1	f3-6mer
27	adga YGCSGS IYELSGPLC ^e	2	f3-15mer
28	aENPMYSAA	1	f8-8mer
29	a RSP LQcADcSPPHP	1	f88-Cys2 ^f
30	a QTWEKc GVDcRNLQ	1	f88-Cys3
31	a STNMc HPYDcDQLP	1	f88-Cys4
32	a EFStc YRDPmCHKPG	1	f88-Cys5
33	a MFTPc GTDDAcEGST	1	f88-Cys5
34	a EFMTc G STGR EcGVNT ^g	1	f88-Cys6
35	a TMSLc DPSKTWcSATN	1	f88-Cys6

^a The sequences of the displayed peptides are given from the N terminus of the mature coat protein (just downstream of the signal peptidase cleavage site) through the last randomized amino acid (see also footnote c). Randomized amino acids are shown in capital letters; amino acids that are constant in the given library are shown in lowercase letters. Sequence motifs are bold, as are peptides that match segments of T4 proteins (footnotes e and g).

^b Number of clones with the indicated structure among the 43 sequenced clones.

^c The bracketed peptide was also fused to carriers IPIII and MBP.

^d This peptide has the correct length for the f88-Cys0 library but has a Tyr-Arg dipeptide in place of the fixed Cys-Cys dipeptide.

^e The highlighted IYELS peptide matches positions 117–121 of T4 protein gp15.

^f This clone ostensibly derives from library f88-Cys1, but its sequence indicates that it is a contaminant from library f88-Cys2.

^g The highlighted GSTGR peptide matches positions 114–118 of T4 protein gp36.

absorbed antiserum serially diluted in TTDBA. T4-absorbed and mock-absorbed samples of each serum were assayed on the same dish to allow side-by-side comparison. To correct for dish-to-dish variation, 10 wells on each dish were coated with a primary reference standard (purified phage displaying RAP 5, Table II) and reacted with serial 2-fold dilutions of direct anti-T4 antiserum. After reacting overnight, the wells were washed with TBS/Tween and reacted with alkaline phosphatase-conjugated goat anti-mouse Fc γ (45 ng/ml; 150 μl ; Pierce) for 2 h at room temperature. Dishes were then washed with TBS/Tween and incubated for 1 h with *p*-nitrophenylphosphate chromogenic substrate while the OD was monitored (23). The slope MOD/min was estimated for each well and used as the ELISA signal. Percentage of cross-reactivity for a given serum was calculated as $100 \times (1 - T_{\text{path}}/T_{\text{mock}})$, where T_{mock} is the standardized titer for the mock-absorbed sample of the antiserum and T_{path} is the corresponding standardized titer for the T4-absorbed sample of the antiserum.

Anti-peptide titers of indirect antisera elicited by IPIII-displayed peptides were also measured by ELISAs in which the immobilized Ags were MBP-displayed peptides rather than phage-displayed peptides. For these assays, wells of ELISA dishes were coated with 40 μl of 10 $\mu\text{g}/\text{ml}$ streptavidin in 0.1 M NaHCO₃, washed with TBS/Tween, and reacted with 100 μl of 1 $\mu\text{g}/\text{ml}$ biotinylated fusion protein diluted in TTDBA. After washing with TBS/Tween, the wells were reacted with dilutions of T4-absorbed and

Table III. Natural antigenic peptides affinity selected with direct anti-T4 Ab

NAP	No. of Clones ^a	Peptide Sequence ^b	T4 Positions ^c
gp34 1-97 ^{d,e}	1	atk*epgrM[AAEIKRKFRAEDGLDAGGDKIINVAL ADRTVGTGDNVNDYLIQENTVQQYDPTRGYKDFV IIYDNRFWAAINDIPKPGAFNSGRWRALRTDANWIAae]	150785-151084
Hoc 1-89 ^d	4	a*g*lm[TFTVDITPKTPTGVIDETKQFTATPSGQT GGGTITYAWSVDNVPQDGAEEATFSYVLKGPAGQKT IKVVATNTLSEGGPETAEATTTITaae]	111326-111049
Hoc 320-347	1	aEEGKDWKTDDPDSKYLLHRYTLQKMMKDaae	110357-110275
Wac 461-487 ^f	7	[aPEAPRDGQAYVRKDGWVLLSTFLSPAaae]	93506-93585
Pin 130-145 ^g	2	nvpaaadAAEYIEWLETQISFSDaae	46993-46946
gp9 32-49 1-93 ^h	1	· · · MFIQEPKKLIDTGEIGNASTGDILFDGNGKINSDF FNAIYNAFGDQRKMAVANGTGADGQIIHATGYGQK HSITEYATPVKVGTRHDIDTSTVGaae	<87216-87493
25-49	1	aFDGNGKINSDFNAIYNAFGDQRKMAaae	87288-87361
2-51	1	aFIQEPKKLIDTGEIGNASTGDILFDGNGKINSDF NAIYNAFGDQRKMAVaae	87219-87364
20-49 ⁱ	2	[TGDILFDGNGKINSDFNAIYNAFGDQRKMAaae]	87276-87361
21-51	1	aGDILFDGNGKINSDFNAIYNAFGDQRKMAVaae	87276-87364
24-49	1	aLFDGNGKINSDFNAIYNAFGDQRKMAaae	87285-87361
32-55 ^j	1	nvpaa[NSDFNAIYNAFGDQRKMAVANGTGaae]	87309-87379
4-49	1	aQEPKKLIDTGEIGNASTGDILFDGNGKINSDFNA IYNAFGDQRKMAaae	87225-87361
4-51	1	aQEPKKLIDTGEIGNASTGDILFDGNGKINSDFNA IYNAFGDQRKMAVaae	87225-87367
1-63 ^d	1	apiimamdqteeinilftf*ikgahglpfs1*il* thketaM[FIQEPKKLIDTGEIGN ASTGDILFDGNGKINSDFNAIYNAFGDQRKMAVANGTGA DGQIIHAae]	87096-87395

^a Number of clones with the indicated structure among the 26 sequenced clones.

^b Peptide sequences extend from the N terminus of the mature coat protein (after the signal peptidase cleavage site) through the Ala-Ala-Glu tripeptide, as shown in Fig. 1. Capital letters show the longest continuous stretch of amino acids that match residues of the corresponding T4 protein; other amino acids are written in lowercase letters. Asterisks indicate in-frame stop codons upstream of the initiator Met (footnote d). Square brackets enclose peptides that were fused to IPIII and MBP carriers. The 18-residue sequence that is common to all the gp9 peptides is highlighted.

^c Positions from GenBank accession no. AF158101.3 that correspond to the genomic DNA fragment inserted into the phage display vector.

^d The T4 inserts in these clones include nucleotides upstream of the ATG initiator codon for the indicated T4 gene. In the given peptide sequence, the amino acids (lowercase letters) and stop codons (*) specified by these nucleotides in the T4 gene's reading frame are shown immediately following the N-terminal Ala. The reactivity of these phage with direct anti-T4 antiserum leaves little doubt that the NAP peptides are actually displayed on the phage surface despite the stop codons; such phages have been described previously (73). Nevertheless, we speculate that the copy number of these peptides is particularly low, not only because of the stop codons but also because of their exceptional length.

^e The AAG codon for the second Lys is a GAA Glu codon in GenBank.

^f The TTG codon for the first Leu is a TTC Phe codon in GenBank; the Leu agrees with the published structure of the Wac protein (69).

^g Alternative construct as defined in Fig. 1; the first five nucleotides of the Ala-Asp codons, GCTGA, are of unknown origin.

^h The ellipsis (· · ·) indicates that the available sequence information is incomplete at this end. Because of its exceptional length, this peptide is probably displayed at low copy number on the phage carrier.

ⁱ Linker mutation causes Ala→Thr replacement at the N terminus of mature protein.

^j Alternative construct as defined in Fig. 1.

mock-absorbed indirect antisera and processed as described above. The same primary reference standard was used for anti-peptide titers obtained with either MBP-displayed peptides or phage-displayed peptides.

Anti-T4 titer

Indirect antisera elicited by phage-displayed peptides were reacted with wild-type T4; indirect antisera elicited by IPIII-displayed peptides were reacted with a mutant form of T4 lacking the IPIII protein, thus avoiding interference by anti-IPIII Abs elicited by the IPIII carrier. Wells of ELISA dishes were coated with 5×10^9 wild-type or IPIII-less T4 particles in 50 μ l D-PBS, washed with TBS/Tween, and reacted overnight with 100 μ l indirect antisera diluted in TTDBA. Because antisera elicited by phage-displayed peptides had weak anti-T4 activity, anti-T4 reactivities of these antisera were determined at a single 1/10 dilution and precise titers could not be determined. Indirect antisera elicited by IPIII-displayed peptides were serially diluted and assayed by ELISA as described in the previous subsection. Titers were compared with the same primary reference standard to which all other titers were referred.

Results

Selection of antigenic peptides

Two distinct bacteriophage were used in this work: the filamentous phage that are the carriers of the peptides in the NPL and RPLs, and the T4 phage that serve as the model pathogen. To avoid

confusion in what follows, we will reserve the term phage (and all related virological terms) for the filamentous phage carriers, referring to T4 as the pathogen, the T4 pathogen, or simply T4.

The direct Ab used for affinity selection in this project was the IgG fraction of serum from mice that had been hyperimmunized with T4. The total IgG population, as well as the serum it derives from, will be referred to informally as direct anti-T4 Ab, even though only a fraction of the component molecules are actually specific for T4 epitopes.

The phage display libraries that served as a source of antigenic peptides are listed in Table I. Display of guest peptides on the surface of the phage particles in the libraries is achieved by splicing a short DNA coding sequence into the gene for a host phage coat protein, thus genetically fusing the guest peptide to the host polypeptide. Eleven of the phage display libraries are RPLs, in which the guest coding sequences are degenerate synthetic oligonucleotides. The RPLs differ with respect to the number of randomized amino acid residues, the structural constraints imposed on the displayed peptide, the host coat protein, and the number of peptides displayed per virion (Table I).

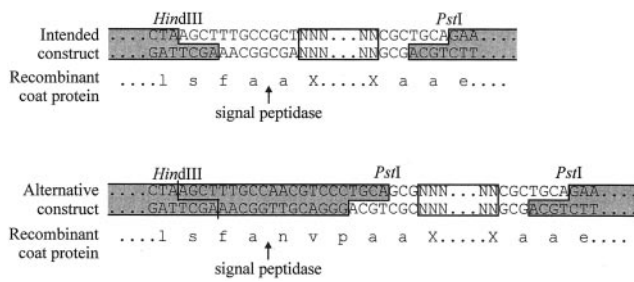


FIGURE 1. Structure of the NPL. Ligation of hairpin linkers to blunt-ended genomic fragments (boxed in the diagram) results in double-stranded fragments whose strands are connected at both ends by loops; DNA that does not have this intended structure is degraded by successive digestion with exonucleases III and VII. Surviving molecules are digested with *HindIII* and *PstI*, gel-purified, and spliced to f88-4 vector that has been cleaved at its *HindIII* and *PstI* sites and freed of the short stuffer that lies between. The intended construct that results from this process is shown, along with the recombinant coat protein it encodes. Also shown at the bottom is an alternative construct that was found in some NAP clones. This construct carries a genomic fragment with *PstI* hairpin linkers ligated to both ends; after cleavage with *PstI*, such a fragment could be spliced into an f88-4 vector molecule whose *HindIII* site remained uncut during the restriction digestion. In both constructs, nucleotides derived from the f88-4 vector are shaded.

The twelfth library is an NPL, in which the guest coding sequences are random fragments of genomic DNA from the model pathogen (Table I); Fig. 1 outlines its structure. In such NPLs, some of the phage clones display fragments (~20–100 amino acids) of actual T4 pathogen polypeptides. Successful natural peptide display requires that the genomic insert encode part of a structural component of T4 and that it be spliced into the vector so that its natural reading frame is correctly fused to that of the host coat protein gene. Because of the randomness of the genomic inserts, only a minority of the clones in the NPL meet these requirements; the remainder display no guest peptide at all, display part of a T4-encoded protein that is not present on T4 particles, or display a random peptide encoded by a non-natural reading frame. Despite their relative scarcity, clones displaying natural peptide fragments of pathogen proteins might be a rich source of peptide epitopes that mimic native antigenic determinants.

Most affinity-selected RAPs were mimotopes

Direct anti-T4 Ab was used to affinity select RAPs from the 12 phage display libraries. A majority of the RAP sequences could be grouped into one of two prominent motif families having consensus sequences EWxPPxR (RAPs 1–7, Table II) or EFPY (RAPs 8–21). Both motifs were selected from several of the 11 RPLs. Two minor motifs, FWWGY (RAPs 22–23) and EMNYxxxS (RAPs 24–25), were represented by a few clones each, and 10 RAPs (26–35) could not be grouped into clear motif families. RAPs 27 and 34 are the only two RAPs that potentially align with segments of T4 polypeptides (bold residues in Table II). The remaining RAPs or RAP motifs are mimotopes (49, 50).

NAPs represent six natural T4 Ags

Table III lists the T4 protein fragments represented among the gene product affinity-selected NAPs. Five of the NAPs correspond to defined segments of T4 proteins Wac, Hoc (two separate segments), gene product gp34, and gp9. The sixth corresponds to part of protein Pin, which is encoded by the T4 genome but is not present in the T4 particle. The direct Ab used for affinity selection presumably does not include subspecificities induced by Pin itself. In effect, then, the mimicking segment on Pin is another RAP.

The gp9 epitope stands apart from the other NAPs in that it is actually a family of overlapping peptides with a common core spanning gp9 residues 32–49 (bold residues in Table III). Evidently this epitope is non-context dependent, maintaining its binding activity in many different contexts of flanking amino acids. The Wac, Hoc, and gp34 epitopes, in contrast, are arguably context dependent, because all clones carrying one of these epitopes are identical. In general, non-context-dependent epitopes such as gp9 32–49 will be far more abundant in the original NPL than context-dependent epitopes. Perhaps this is the reason that the gp9 epitope was the most abundantly represented among the selected NAPs.

All RAPs mimic T4 epitopes

Because the IgG used for affinity selection undoubtedly contained many background subspecificities against Ags other than T4, it could not be assumed in advance that the selected peptides correspond to T4 epitopes. In the context of a real disease, extensive screening for reactivity with positive and negative sera is used to identify peptides that are likely to correspond to disease-related natural epitopes (21, 51, 52). However, because of the availability of purified T4 as an immunochemical reagent in our model system, pathogen specificity was assessed by an easier, more direct technique: the direct anti-T4 IgG was depleted of T4-specific Abs by immunoabsorption, and the anti-peptide reactivity of the T4-absorbed IgG was compared with that of mock-absorbed IgG using ELISA. All RAPs exhibited greatly reduced ELISA reactivity when assayed with T4-absorbed as compared with mock-absorbed IgG (data not shown), implying that they bind the same direct-Ab subspecificities as T4. The same was true of the Hoc 1–89, Wac 461–487, gp34 1–97, and Pin 130–145 NAPs (data not shown). In the case of the Pin epitope, this result strengthens the supposition above that the direct Ab subspecificities selecting this peptide are directed against an epitope that is present on T4 but that is mimicked by Pin residues 130–145. Thus, despite the fact that all but two of the RAPs lack any similarity to actual T4 sequences, all appear to bind specifically to T4-induced subspecificities. This conclusion was corroborated by the finding that the direct antiserum had high titers against all the tested RAPs and NAPs, while the corresponding preimmune serum had little or none (data not shown). Surprisingly, T4 absorption did not remove a significant fraction of the titers to the Hoc 320–347 or gp9 32–49 NAPs; possibly, the direct Ab subspecificities that affinity selected these two peptides were elicited by denatured epitopes that are present in the T4-immunized mice but are absent or scarce when T4 is used as an immunoabsorbent.

Assessing immunogenic fitness

In an initial, large-scale survey, immunogenic fitness was evaluated for 13 representative RAPs and eight representative NAPs. For this survey, the affinity-selected phage were themselves used as carriers for immunization. Using peptide-bearing phage directly as immunogens (53, 54) avoids the need to transfer the peptides to an alternative immunogenic carrier. The resulting indirect antisera were titered by ELISA against the T4 pathogen and the peptide. This strategy was convenient for the large-scale survey, but a few peptides could not be analyzed because they did not provoke an adequate response. Therefore, a subset of the antigenic peptides was displayed on alternative carriers and reinvestigated, as will be described.

Most phage-displayed peptides induced strong anti-peptide titers

Mice were hyperimmunized with the phage-displayed peptides, and the anti-peptide titers of the resulting indirect antisera were measured by ELISA after removing Abs against the phage carrier

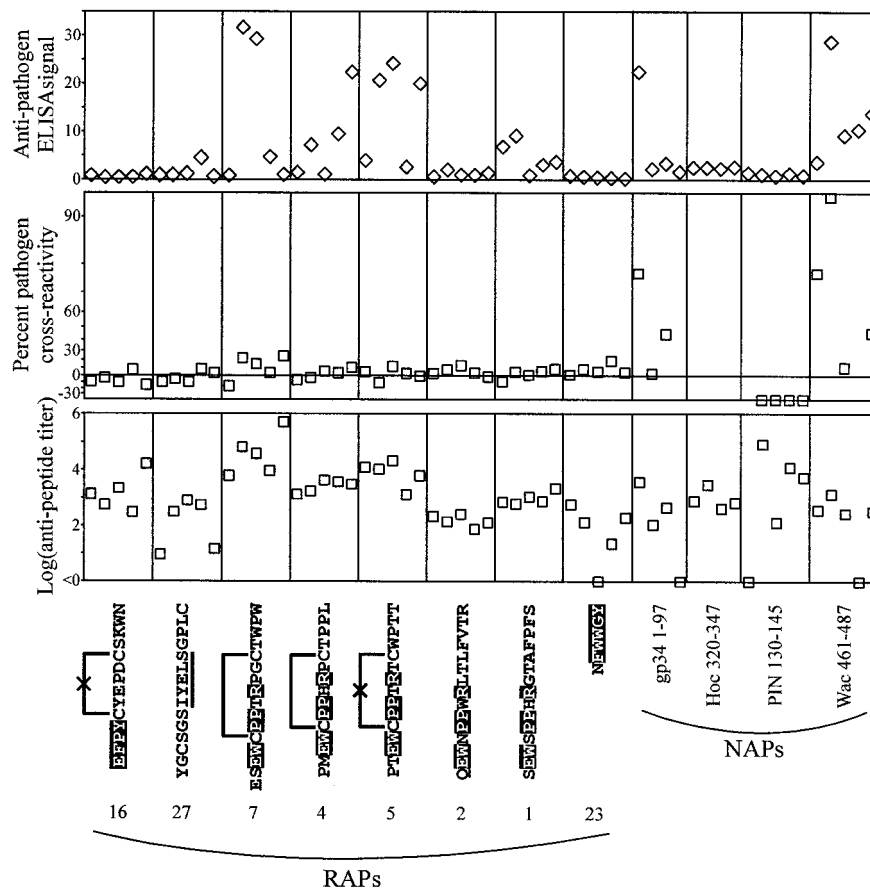


FIGURE 2. Anti-T4 signal, percentage of cross-reactivity, and anti-peptide titer of indirect antisera elicited by immunization with phage-displayed peptides. Each peptide-bearing phage was used to immunize five mice (horizontal panels with peptides labeled on the x-axis). Numbers below RAP sequences indicate RAP number as given in Table II. RAP motifs (bold residues in Table II) are highlighted. The underlined portion of the RAP 27 peptide matches positions 117–121 of T4 protein gp15 (Table II). Brackets on the RAP peptide sequences at the *bottom* indicate disulfide bonds; the disulfide bonds marked with **X** are not required for binding to direct Ab, whereas the others are required (data not shown). Indirect antiserum from each mouse was assayed as indicated, with results for each indirect antiserum aligned vertically. \diamond , Results from ELISAs in which T4 served as immobilized Ag; \square , results from ELISAs in which the corresponding phage-displayed peptide served as the immobilized Ag. *Upper panel*, Anti-T4 reactivity measured by ELISA against immobilized T4 at a single serum dilution of 1/10; *middle panel*, percentage of cross-reactivity measured as described; *lower panel*, anti-peptide titer measured as described. The following RAPs lacked detectable anti-T4 reactivity despite adequate anti-peptide titers (data not shown): 9, 12, 13, 24, 34 (Table II).

by absorption. Fig. 2, *lower panel*, shows results from a representative sample of mice; data points for a single indirect antiserum are aligned vertically in Fig. 2. Most mice responded strongly to the phage-displayed RAPs, generating anti-peptide titers on the order of 10^3 – 10^6 ; as usual with anti-peptide responses, there was some mouse-to-mouse variation in the response to an individual peptide (21, 51). Four of the NAPs (gp34 1–97, Pin 130–145, Hoc 320–347, and Wac 461–487) elicited anti-peptide titers on the order of 10^2 – 10^4 . NAPs Hoc 1–89, gp9 1–63, gp9 20–49, and gp9 32–55 failed to elicit any detectable anti-peptide response (data not shown). The NAPs were thus generally weaker immunogens than the RAPs when administered on their original phage carrier.

Several RAPs and NAPs elicited antisera that react with T4

The anti-T4 reactivity of the indirect antisera was measured by ELISA at a single serum dilution of 1/10. Antisera elicited by four of the five tested RAPs in the EWxPPxR family, and by the Wac 461–487 and gp34 1–97 NAPs, reacted measurably with T4, indicating some degree of immunogenic fitness (Fig. 2, *upper panel*). In contrast, indirect antisera elicited by the Hoc 320–347 and Pin 130–145 NAPs, and by all RAPs in the EFPY family, lacked any detectable anti-T4 reactivity, even though they had adequate anti-

peptide titers (Fig. 2, *upper panel*). Immunogenic fitness could not be meaningfully assessed for the NAPs that failed to elicit a detectable anti-peptide response.

Because many of the indirect antisera showed weak or undetectable anti-T4 reactivity, anti-T4 reactivity could not be quantified by the more accurate method of measuring signals at a series of serum dilutions. Even if anti-T4 reactivity could be measured this way, it would provide an imperfect assessment of immunogenic fitness because it is complicated by two confounding factors: immunogenicity and the copy number of the cognate native epitope on T4 (which serves as immobilized Ag in the ELISA). Thus, although the qualitative data reported in this work suffice to demonstrate a degree of immunogenic fitness on the part of some antigenic peptides, we sought to assess immunogenic fitness more quantitatively.

Using the percentage of cross-reactivity to assess immunogenic fitness

We assessed immunogenic fitness quantitatively by determining the percentage of indirect Abs that cross-react with the T4 pathogen (percentage of cross-reactivity). To measure the percentage of cross-reactivity, intact T4 was used as an immunoabsorbent to

completely deplete each indirect antiserum of all detectable T4-reactive Abs; as a control, matched samples of antisera were mock-absorbed by carrying out exactly the same steps in parallel without T4 (or, in the case of the Wac and Hoc NAPs, with a mutant form of T4 missing the Wac or Hoc protein). The T4-absorbed and mock-absorbed antisera were then titered side-by-side against the peptide. The results in Fig. 2, *middle panel*, are plotted linearly in terms of $\log(T_{\text{mock}}/T_{\text{path}})$, where T_{mock} and T_{path} are the mock-absorbed and T4 pathogen-absorbed titers, respectively; experimental uncertainties are expected to be roughly constant over this scale. The nonlinear scale on the ordinate axis in Fig. 2 gives the equivalent values of the percentage of cross-reactivity calculated as $100 \times (T_{\text{mock}} - T_{\text{path}})/T_{\text{mock}}$. Unlike overall anti-pathogen titer, the percentage of cross-reactivity is independent of both immunogenicity and copy number of the cognate epitope on the pathogen (assuming pathogen absorption is complete). It is much less sensitive to weak immunogenic fitness than is overall anti-pathogen reactivity, because it is proportional to the relatively small difference between two relatively large numbers, T_{mock} and T_{path} . Therefore, only peptides with superior immunogenic fitness will pass this stringent test.

None of the indirect antisera elicited by any of the RAPs showed a significant percentage of cross-reactivity, as is shown for a representative sampling in Fig. 2, *middle panel*. This was true even of antisera with detectable overall anti-T4 reactivity; evidently the T4-reactive Abs in those antisera comprise only a small fraction of the total anti-peptide response. The RAPs tested in this way include five in the EWxPPxR family that differ with respect to cysteine bridges (Fig. 2); the data thus do not support the hypothesis that immunogenic fitness can be dramatically enhanced by the simple expedient of installing fixed disulfide constraints (11, 15, 55, 56). Neither RAP that potentially aligns with segments of T4 polypeptides passed this stringent test of immunogenic fitness (RAP 27, Fig. 2; RAP 34, data not shown).

Indirect antisera elicited by the Pin 130–145 and Hoc 320–347 NAPs also lacked detectable percentages of cross-reactivity, in accord with their lack of detectable anti-T4 reactivity. In contrast, several indirect antisera induced by the Wac 461–487 and gp34 1–97 NAPs showed substantial percentages of cross-reactivity.

High anti-peptide titers were achieved when IPIII fusion proteins served as immunogens

Immunogenic fitness could not be assessed for a few of the peptides in the initial survey because they failed to elicit an adequate indirect Ab response. A plausible reason for failure was low-density display on the phage carrier. Low display density is particularly likely in the case of the gp34 1–97, Hoc 1–89, and gp9 1–63 NAPs, which are exceptionally long and/or are encoded by inserts with in-frame stop codons (Table III). When used as an immunogen, a phage-borne peptide displayed at low copy number may elicit a poor yield of indirect Abs, regardless of intrinsic immunogenicity. Furthermore, when used subsequently as the immobilized ELISA Ag, it presents fewer target ligands for Abs to bind, thus reducing the ELISA signal at a given concentration of anti-peptide Ab.

Accordingly, a subset of antigenic peptides were fused to two unrelated carrier proteins: IPIII (an internal protein of T4) as the carrier for immunization and MBP of *E. coli* as the carrier for ELISA Ags. The Pin 130–145 and Hoc 320–347 NAPs and all RAPs outside the EWxPPxR motif family were excluded from this reinvestigation because they showed no hint of immunogenic fitness in the initial survey despite provoking adequate anti-peptide titers.

Each IPIII-displayed peptide was used to hyperimmunize 5–15 mice; the resulting indirect antisera were titered by ELISA against the immunizing peptide, using both phage and MBP fusion proteins as the immobilized ELISA Ags. Results are graphed in Fig. 3, *lower panel*; \square and \times indicate data for the phage-displayed and MBP-displayed ELISA Ags, respectively. Data for a single antiserum, including multiple independent repetitions of each measurement, are aligned vertically in Fig. 3.

As shown in Fig. 3, *lower panel*, antisera induced by four of the IPIII-displayed peptides—RAP 5, Wac 461–487, gp34 1–97, and Hoc 1–89—had strong titers (roughly 10^5) against MBP-displayed ELISA Ags. Thus, when displayed at equal densities on a defined immunogen and arrayed at equal densities on the ELISA wells, these peptides have comparable immunogenicities. Antisera to the peptides that are presumably poorly displayed on phage (gp34 1–97 and Hoc 1–89) had much lower titers against phage than against MBP fusion proteins, while antisera to peptides that are presumably well displayed on phage (RAP 5 and Wac 461–487) gave slightly higher titers against phage than against MBP fusion proteins.

None of the gp9 peptides tested elicited usable anti-peptide titers, regardless of whether phage or IPIII served as the carrier for immunization, or whether phage-displayed or MBP-displayed peptides served as the immobilized ELISA Ags (data not shown). Therefore, these peptides seem to be intrinsically poor immunogens.

Three of six NAPs had exceptional immunogenic fitness

Fig. 3, *middle panel*, shows the percentages of cross-reactivity for indirect antisera elicited by the IPIII-displayed peptides. Figures

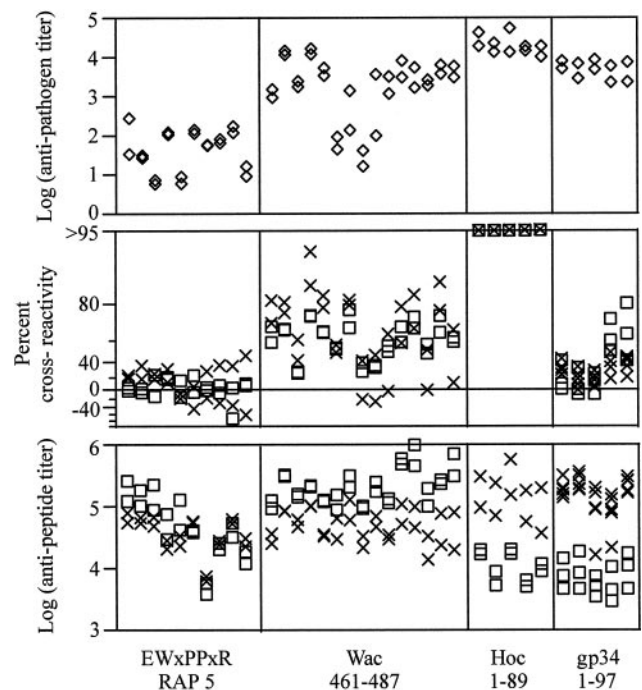


FIGURE 3. Anti-T4 titer, percentage of cross-reactivity, and anti-peptide titer of indirect antisera elicited by IPIII-displayed peptides. Each peptide-IPIII fusion protein was used to immunize 5–15 mice (horizontal panels with peptides labeled on the x-axis). Indirect antiserum from each mouse was assayed as indicated, with results for each indirect antiserum aligned vertically. \diamond , Results from ELISAs in which T4 served as Ag; \square , results from ELISAs in which the corresponding phage-borne peptide served as Ag; \times , results from ELISAs in which the corresponding peptide MBP-fusion protein served as the Ag. *Upper panel*, Anti-T4 titer measured as described; *middle panel*, percentage of cross-reactivity measured as described; *lower panel*, anti-peptide titer measured as described.

for a given antiserum were similar whether measured using phage or MBP fusion proteins as the ELISA Ag, even when the corresponding anti-peptide titers were markedly different. This was expected, because the percentage of cross-reactivity depends on the ratio of pathogen-absorbed vs mock-absorbed anti-peptide titers, not their absolute values.

The results for RAP 5 and the Wac 461–487 and gp34 1–97 NAPs confirm and extend the previous results with antisera elicited by phage-displayed peptides. Antisera elicited by RAP 5 had no measurable cross-reactivity, while antisera elicited by the Wac 461–487 and gp34 1–97 NAPs had cross-reactivities of up to 93 and 80%, respectively. The results also reveal that the Hoc 1–89 NAP is even more immunogenically fit than the other NAPs, inducing indirect antisera with 100% cross-reactivity.

The anti-T4 titers of the indirect antisera are graphed in Fig. 3, *upper panel*, and support the conclusions from the percentage of cross-reactivity measurements. The highest anti-T4 titers were elicited by the Hoc 1–89 NAP, which shows maximal immunogenic fitness as reported above, and which has a high copy number (160) on T4. The anti-T4 titers elicited by the Wac 461–487 and gp34 1–97 NAPs were only ~10- to 100-fold lower, even though their cognate native epitopes have much lower copy numbers (6) on T4. Finally, the EWxPPxR RAP 5 provoked much lower anti-T4 titers, in accord with the very low percentages of cross-reactivity of these sera (the copy number of the as-yet-unidentified native epitope mimicked by this RAP is unknown but is almost certainly at least 6).

In summary, antisera against the Hoc 1–89, Wac 461–487, and gp34 1–97 NAPs had uniformly high anti-peptide titers, good to outstanding percentages of cross-reactivity, and excellent overall anti-pathogen reactivities. On the score of both immunogenicity and immunogenic fitness, at least, any of these three peptides would be an excellent candidate for synthetic vaccine development. By the same criteria, the most successful of the RAPs would make a considerably less attractive candidate, although even this peptide is highly immunogenic and shows some degree of immunogenic fitness.

Discussion

This paper reports a critical appraisal of antigenic peptides for immunogenic fitness, using a model pathogen that allows immunogenic fitness to be quantified. Antigenic peptides were obtained through epitope discovery, a high throughput process that should be feasible in the context of almost any infectious disease regardless of the idiosyncratic details of its pathogenesis.

As in previous epitope discovery projects, we identified RAPs that are immunogenically fit according to the criterion that they induce indirect antisera that cross-react with the original pathogen (2, 11, 21, 51, 57–64). However, by the more stringent criterion of percentage of cross-reactivity, even the most successful of the RAPs in this study achieved only marginal immunogenic fitness. In contrast, three of the six NAPs tested turned out to have excellent immunogenic fitness, eliciting indirect antisera with high anti-peptide titers, high anti-pathogen titers, and a substantial percentage of cross-reactivity. Although there is no equivalent systematic study of immunogenic fitness with which it can be compared, this 50% success rate is almost certainly much higher than those achieved historically.

Because immunogenic fitness has rarely been measured in isolation, the exact relationship between efficacy and immunogenic fitness is largely unmapped. Nevertheless, there are good reasons to think that immunogenic fitness is strongly correlated with efficacy. By definition, a peptide with poor immunogenic fitness elicits a preponderance of indirect Abs that do not cross-react with the

pathogen and therefore do not contribute to protection. It is unlikely that these nonprotective indirect Abs would directly interfere with protection. However, by dominating Ab production at the expense of pathogen cross-reacting specificities, they could limit the protective titer that can be achieved by vaccination. Our results are consistent with this effect. Hyperimmunization with RAP 5 (EWxPPxR) and with the three best NAPs gave comparable overall anti-peptide titers (Fig. 3, *lower panel*). However, the immunogenically fit NAPs elicited much higher anti-pathogen titers than did RAP 5, which had marginal immunogenic fitness (Fig. 3, *upper panel*).

The superior immunogenic fitness of the NAPs selected in this work may result from extensive geometric mimicry of the corresponding natural epitopes, including not only the binding valences that actually contact Ab but also the surrounding structure that holds those valences in a conformation favorable for binding. Their relatively large size (27–97 amino acids for the three most immunogenically fit NAPs) provides ample opportunities for reproducing multiple conformation-stabilizing interactions present in the intact native pathogen (9, 12). In addition, these self-folding native-like domains may be large enough to encompass more than one epitope (33). In contrast to an NPL, an RPL probably contains few large ensembles of amino acids able to closely mimic the structure of a native pathogen. Unless the RPL includes an impossibly huge number of peptides, mimicry on the part of a RAP will seldom extend beyond a small handful of critical binding residues, even if the overall length of the RAPs in the RPL is much longer than the 6–17 amino acids of the RAPs studied in this work.

The degree of mimicry afforded by peptides affinity selected from RPLs frequently translates into sufficient immunogenic fitness to induce indirect antisera that cross-react with the original pathogen. There have been several studies in which RAPs have apparently achieved sufficient immunogenic fitness to provide some measure of disease protection (2, 8, 57, 64, 65), although other studies have been less promising (7, 18). Furthermore, there are undoubtedly some discontinuous native epitopes that cannot be reproduced by fragments of the corresponding proteins, and which therefore can only be mimicked by RAPs (2, 11, 66, 67). Nevertheless, our results suggest that NPLs may be a superior source of peptides with exceptionally good immunogenic fitness and therefore provide an important alternative to RPLs. Undoubtedly, only a tiny minority of the displayed peptides in an NPL contain self-folding subdomains that closely resemble native structures, but those peptides might be greatly favored during affinity selection. Although the RPLs are at least as diverse as the NPL, they are presumably a far poorer source of such native-like structural subdomains, which might resemble the small pathogen polypeptides that have succeeded as subunit vaccines (e.g., tetanus toxoid or the recombinant hepatitis B surface Ag).

The immunogenically fit Wac 461–487 NAP appears to be just the sort of self-folding subdomain envisioned in the previous paragraphs. The Wac protein has been subjected to detailed structural analysis. The bulk of the protein consists of a long trimeric coiled coil domain, with properties similar to those of other fibrous proteins such as collagen (68, 69). The Wac NAP corresponds exactly to a trimeric, globular domain at the C-terminal end of the fibrous coiled coil (70). The structural features of this domain suggest that it can maintain its conformation outside the context of the intact Wac protein.

Detailed structural information for pathogen proteins is unlikely to be available in a vaccine development project. However, it may be possible to determine whether a NAP has any strongly preferred three-dimensional structure, using circular dichroism or nuclear

magnetic resonance. Surveying NAPs for the presence of a preferred structure could help to narrow the search for peptides that are particularly likely to be immunogenically fit.

Admittedly, there is more to vaccine efficacy than immunogenic fitness. There are examples of epitopes or Ags that elicit a strong anti-pathogen Ab response without actually protecting against the disease. Moreover, B epitopes, whether obtained through epitope discovery or conventional Ag dissection, must usually be combined with a source of appropriate Th epitopes to fashion an effective vaccine. Nevertheless, establishment of a simple, generic process for discovering immunogenically fit peptides should significantly advance efforts to develop synthetic vaccines.

Acknowledgments

We thank John R. Marston and the University of Missouri Department of Laboratory Animal Medicine for technical assistance.

References

- Arnon, R., and R. J. Horwitz. 1992. Synthetic peptides as vaccines. *Curr. Opin. Immunol.* 4:449.
- Steward, M. W., C. M. Stanley, and O. E. Obeid. 1995. A mimotope from a solid-phase peptide library induces a measles virus-neutralizing and protective antibody response. *J. Virol.* 69:7668.
- Garrity, R. R., G. Rimmelzwaan, A. Minassian, W. P. Tsai, G. Lin, J. J. de Jong, J. Goudsmit, and P. L. Nara. 1997. Refocusing neutralizing antibody response by targeted dampening of an immunodominant epitope. *J. Immunol.* 159:279.
- Taboga, O., C. Tami, E. Carrillo, J. I. Nunez, A. Rodriguez, J. C. Saiz, E. Blanco, M. L. Valero, X. Roig, J. A. Camarero, et al. 1997. A large-scale evaluation of peptide vaccines against foot-and-mouth disease: lack of solid protection in cattle and isolation of escape mutants. *J. Virol.* 71:2606.
- Zanetti, M., E. Sercarz, and J. Salk. 1987. The immunology of new generation vaccines. *Immunol. Today* 8:18.
- Ben-Yedidia, T., and R. Arnon. 1997. Design of peptide and polypeptide vaccines. *Curr. Opin. Biotechnol.* 8:442.
- El Kasmi, K. C., S. Deroo, D. M. Theisen, N. H. Brons, and C. P. Muller. 1999. Crossreactivity of mimotopes and peptide homologues of a sequential epitope with monoclonal antibody does not predict crossreactive immunogenicity. *Vaccine* 18:284.
- Grabowska, A. M., R. Jennings, P. Laing, M. Darsley, C. L. Jameson, L. Swift, and W. L. Irving. 2000. Immunisation with phage displaying peptides representing single epitopes of the glycoprotein G can give rise to partial protective immunity to HSV-2. *Virology* 269:47.
- Kaumaya, P. T., A. M. VanBuskirk, E. Goldberg, and S. K. Pierce. 1992. Design and immunological properties of topographic immunogenic determinants of a protein antigen (LDH-C4) as vaccines. *J. Biol. Chem.* 267:6338.
- Van Regenmortel, M. H. V. 1989. Structural and functional approaches to the study of protein antigenicity. *Immunol. Today* 10:266.
- Lundin, K., A. Samuelsson, M. Jansson, J. Hinkula, B. Wahren, H. Wigzell, and M. A. Persson. 1996. Peptides isolated from random peptide libraries on phage elicit a neutralizing anti-HIV-1 response: analysis of immunological mimicry. *Immunology* 89:579.
- Shapira, M., M. Jibson, G. Muller, and R. Arnon. 1984. Immunity and protection against influenza virus by synthetic peptide corresponding to antigenic sites of hemagglutinin. *Proc. Natl. Acad. Sci. USA* 81:2461.
- Jackson, D. C., J. M. Murray, D. O. White, C. N. Fagan, and G. W. Tregear. 1982. Antigenic activity of a synthetic peptide comprising the "loop" region of influenza virus hemagglutinin. *Virology* 120:273.
- Zhong, G., I. Toth, R. Reid, and R. C. Brunham. 1993. Immunogenicity evaluation of a lipidic amino acid-based synthetic peptide vaccine for *Chlamydia trachomatis*. *J. Immunol.* 151:3728.
- Zhong, G., G. P. Smith, J. Berry, and R. C. Brunham. 1994. Conformational mimicry of a chlamydial neutralization epitope on filamentous phage. *J. Biol. Chem.* 269:24183.
- Schulze-Gahmen, U., H. D. Klenk, and K. Beyreuther. 1986. Immunogenicity of loop-structured short synthetic peptides mimicking the antigenic site A of influenza virus hemagglutinin. *Eur. J. Biochem.* 159:283.
- Van Regenmortel, M. H. 2001. Pitfalls of reductionism in the design of peptide-based vaccines. *Vaccine* 19:2369.
- Felici, F., A. Luzzago, A. Folgori, and R. Cortese. 1993. Mimicking of discontinuous epitopes by phage-displayed peptides. II. Selection of clones recognized by a protective monoclonal antibody against the *Bordetella pertussis* toxin from phage peptide libraries. *Gene* 128:21.
- Nestorowicz, A., G. W. Tregear, C. N. Southwell, J. Martyn, J. M. Murray, D. O. White, and D. C. Jackson. 1985. Antibodies elicited by influenza virus hemagglutinin fail to bind to synthetic peptides representing putative antigenic sites. *Mol. Immunol.* 22:145.
- Smith, G. P., and V. A. Petrenko. 1997. Phage display. *Chem. Rev.* 97:391.
- Folgori, A., R. Tafi, A. Meola, F. Felici, G. Galfre, R. Cortese, P. Monaci, and A. Nicosia. 1994. A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *EMBO J.* 13:2236.
- Smith, G. P., and J. K. Scott. 1993. Libraries of peptides and proteins displayed on filamentous phage. *Methods Enzymol.* 217:228.
- Yu, J., and G. Smith. 1996. Affinity maturation of phage-displayed peptide ligands. *Methods Enzymol.* 267:3.
- Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. 1990. Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. USA* 87:6378.
- Devlin, J. J., L. C. Panganiban, and P. E. Devlin. 1990. Random peptide libraries: a source of specific protein binding molecules. *Science* 249:404.
- Parmley, S. F., and G. P. Smith. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* 73:305.
- Scott, J. K., and G. P. Smith. 1990. Searching for peptide ligands with an epitope library. *Science* 249:386.
- Smith, G. P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228:1315.
- Fack, F., B. Hugle-Dorr, D. Song, I. Queitsch, G. Petersen, and E. K. Bautz. 1997. Epitope mapping by phage display: random versus gene-fragment libraries. *J. Immunol. Methods* 206:43.
- Wang, L. F., D. H. Du Plessis, J. R. White, A. D. Hyatt, and B. T. Eaton. 1995. Use of a gene-targeted phage display random epitope library to map an antigenic determinant on the bluetongue virus outer capsid protein VP5. *J. Immunol. Methods* 178:1.
- Holzem, A., J. M. Nahring, and R. Fischer. 2001. Rapid identification of a tobacco mosaic virus epitope by using a coat protein gene-fragment-pVIII fusion library. *J. Gen. Virol.* 82:9.
- Gupta, S., K. Arora, A. Sampath, S. Khurana, S. S. Singh, A. Gupta, and V. K. Chaudhary. 1999. Simplified gene-fragment phage display system for epitope mapping. *BioTechniques* 27:328.
- Perlaiza, B. L., J. P. Sauzet, A. T. Balde, K. Brahimi, A. Tall, G. Corradin, and P. Druilhe. 2001. Long synthetic peptides encompassing the *Plasmodium falciparum* LSA3 are the target of human B and T cells and are potent inducers of B helper, T helper and cytolytic T cell responses in mice. *Eur. J. Immunol.* 31:2200.
- Dulbecco, R., and M. Vogt. 1954. Plaque formation and isolation of pure lines with poliomyelitis viruses. *J. Exp. Med.* 99:167.
- Meissner, P. S., W. P. Sisk, and M. L. Berman. 1987. Bacteriophage λ cloning system for the construction of directional cDNA libraries. *Proc. Natl. Acad. Sci. USA* 84:4171.
- Wickner, W. 1975. Asymmetric orientation of a phage coat protein in cytoplasmic membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 72:4749.
- Conley, M. P., and W. B. Wood. 1975. Bacteriophage T4 whiskers: a rudimentary environment-sensing device. *Proc. Natl. Acad. Sci. USA* 72:3701.
- Ishii, T., and M. Yanagida. 1977. The two dispensable structural proteins (soc and hoc) of the T4 phage capsid: their purification and properties, isolation and characterization of the defective mutants, and their binding with the defective heads in vitro. *J. Mol. Biol.* 109:487.
- Wilson, J. H., J. S. Kim, and J. N. Abelson. 1972. Bacteriophage T4 transfer RNA. III. Clustering of the genes for the T4 transfer RNA's. *J. Mol. Biol.* 71:547.
- Hong, Y. R., and L. W. Black. 1993. An expression-packaging-processing vector which selects and maintains 7-kb DNA inserts in the blue T4 phage genome. *Gene* 136:193.
- Kim, J. S., and N. Davidson. 1974. Electron microscope heteroduplex study of sequence relations of T2, T4, and T6 bacteriophage DNAs. *Virology* 57:93.
- Carlson, K., and E. Miller. 1994. General procedures. In *Molecular Biology of Bacteriophage T4*. J. Karam, ed. Am. Soc. Microbiol., Washington, DC, p. 427.
- Emrich, J. 1968. Lysis of T4-infected bacteria in the absence of lysozyme. *Virology* 35:158.
- Doermann, A. H., F. A. Eiserling, and L. Boehner. 1973. Genetic control of capsid length in bacteriophage T4. I. Isolation and preliminary description of four new mutants. *J. Virol.* 12:374.
- Anderson, S. 1981. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* 9:3015.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY.
- Galfre, G., P. Monaci, A. Nicosia, A. Luzzago, F. Felici, and R. Cortese. 1996. Immunization with phage-displayed mimotopes. *Methods Enzymol.* 267:109.
- Edelhoch, H. 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* 6:1948.
- Geysen, H. M., S. J. Rodda, and T. J. Mason. 1986. A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol. Immunol.* 23:709.
- Geysen, H. M., S. J. Rodda, and T. J. Mason. 1986. The delineation of peptides able to mimic assembled epitopes. *Ciba Found. Symp.* 119:130.
- Prezzi, C., M. Nuzzo, A. Meola, P. Delmastro, G. Galfre, R. Cortese, A. Nicosia, and P. Monaci. 1996. Selection of antigenic and immunogenic mimics of hepatitis C virus using sera from patients. *J. Immunol.* 156:4504.
- Kouzmitcheva, G. A., V. A. Petrenko, and G. P. Smith. 2000. Identifying diagnostic peptides for Lyme disease through epitope discovery. *Clin. Diagn. Lab. Immunol.* 8:150.
- Willis, A. E., R. N. Perham, and D. Wraith. 1993. Immunological properties of foreign peptides in multiple display on a filamentous bacteriophage. *Gene* 128:79.
- Minenkova, O. O., A. A. Ilyichev, G. P. Kishchenko, and V. A. Petrenko. 1993. Design of specific immunogens using filamentous phage as the carrier. *Gene* 128:85.
- Hoess, R. H., A. J. Mack, H. Walton, and T. M. Reilly. 1994. Identification of a structural epitope by using a peptide library displayed on filamentous bacteriophage. *J. Immunol.* 153:724.

56. McConnell, S. J., M. L. Kendall, T. M. Reilly, and R. H. Hoess. 1994. Constrained peptide libraries as a tool for finding mimotopes. *Gene* 151:115.
57. Chargelegue, D., O. E. Obeid, S. C. Hsu, M. D. Shaw, A. N. Denbury, G. Taylor, and M. W. Steward. 1998. A peptide mimic of a protective epitope of respiratory syncytial virus selected from a combinatorial library induces virus-neutralizing antibodies and reduces viral load in vivo. *J. Virol.* 72:2040.
58. Demangel, C., P. Lafaye, and J. C. Mazie. 1996. Reproducing the immune response against the *Plasmodium vivax* merozoite surface protein 1 with mimotopes selected from a phage-displayed peptide library. *Mol. Immunol.* 33:909.
59. Keller, P. M., B. A. Arnold, A. R. Shaw, R. L. Tolman, F. Van Middlesworth, S. Bondy, V. K. Rusiecki, S. Koenig, S. Zolla-Pazner, P. Conard, et al. 1993. Identification of HIV vaccine candidate peptides by screening random phage epitope libraries. *Virology* 193:709.
60. Meola, A., P. Delmastro, P. Monaci, A. Luzzago, A. Nicosia, F. Felici, R. Cortese, and G. Galfre. 1995. Derivation of vaccines from mimotopes: immunologic properties of human hepatitis B virus surface antigen mimotopes displayed on filamentous phage. *J. Immunol.* 154:3162.
61. Motti, C., M. Nuzzo, A. Meola, G. Galfre, F. Felici, R. Cortese, A. Nicosia, and P. Monaci. 1994. Recognition by human sera and immunogenicity of HBsAg mimotopes selected from an M13 phage display library. *Gene* 146:191.
62. Orlandi, R., S. Menard, M. I. Colnaghi, C. M. Boyer, and F. Felici. 1994. Antigenic and immunogenic mimicry of the HER2/neu oncoprotein by phage-displayed peptides. *Eur. J. Immunol.* 24:2868.
63. Stoute, J. A., W. R. Ballou, N. Kolodny, C. D. Deal, R. A. Wirtz, and L. E. Lindler. 1995. Induction of humoral immune response against *Plasmodium falciparum* sporozoites by immunization with a synthetic peptide mimotope whose sequence was derived from screening a filamentous phage epitope library. *Infect. Immun.* 63:934.
64. Arnon, R., R. Tarrab-Hazdai, and M. Steward. 2000. A mimotope peptide-based vaccine against *Schistosoma mansoni*: synthesis and characterization. *Immunology* 101:555.
65. Yu, M. W., J. K. Scott, A. Fournier, and P. J. Talbot. 2000. Characterization of murine coronavirus neutralization epitopes with phage-displayed peptides. *Virology* 271:182.
66. Luzzago, A., F. Felici, A. Tramontano, A. Pessi, and R. Cortese. 1993. Mimicking of discontinuous epitopes by phage-displayed peptides. I. Epitope mapping of human H ferritin using a phage library of constrained peptides. *Gene* 128:51.
67. Barlow, D. J., M. S. Edwards, and J. M. Thornton. 1986. Continuous and discontinuous protein antigenic determinants. *Nature* 322:747.
68. Strelkov, S. V., Y. Tao, M. G. Rossmann, L. P. Kurochkina, M. M. Shneider, and V. V. Mesyanzhinov. 1996. Preliminary crystallographic studies of bacteriophage T4 fibrin confirm a trimeric coiled-coil structure. *Virology* 219:190.
69. Efimov, V. P., I. V. Nepluev, B. N. Sobolev, T. G. Zurabishvili, T. Schulthess, A. Lustig, J. Engel, M. Haener, U. Aebi, S. Venyaminov, et al. 1994. Fibrin encoded by bacteriophage T4 gene *wac* has a parallel triple-stranded α -helical coiled-coil structure. *J. Mol. Biol.* 242:470.
70. Tao, Y., S. V. Strelkov, V. V. Mesyanzhinov, and M. G. Rossmann. 1997. Structure of bacteriophage T4 fibrin: a segmented coiled coil and the role of the C-terminal domain. *Structure* 5:789.
71. Nishi, T., R. J. Budde, J. S. McMurray, N. U. Obeyesekere, N. Safdar, V. A. Levin, and H. Saya. 1996. Tight-binding inhibitory sequences against pp60^{c-src} identified using a random 15-amino-acid peptide library. *FEBS Lett.* 399:237.
72. Petrenko, V. A., G. P. Smith, X. Gong, and T. Quinn. 1996. A library of organic landscapes on filamentous phage. *Protein Eng.* 9:797.
73. Carcamo, J., M. W. Ravera, R. Brissette, O. Dedova, J. R. Beasley, A. Alam-Moghe, C. Wan, A. Blume, and W. Mandrecki. 1998. Unexpected frameshifts from gene to expressed protein in a phage-displayed peptide library. *Proc. Natl. Acad. Sci. USA* 95:11146.