Title
Engineering proteins by phage display

Author
Atwell, Shane,

Publication Date
1999

Peer reviewed|Thesis/dissertation

Engineering Proteins by Phage Display

by

Shane Atwell

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO

Date University Librarian

Degree Conferred: .....

Dedicated to

Sarah Farivar,

confidant,

and

John Galt,

exemplar

Preface

Science involves exchanging values, from the work itself for a bench, a paycheck and mentorship, to reagents and helpful advice, to friendship and simple good will. In the most common and proper case, we know that we will be paid in kind and no sacrifice is involved. These exchanges are voluntary and deserve recognition.

Thank you Jim for being so excited about science, building such a great lab to work in and coaching my writing. Thanks to all the people in Jim's lab who were so friendly and helpful: Dev Sidhu (fellow inventor), Tim Clackson (ditto), Warren DeLano (for help with everything theoretical and computer related), Brian Cunningham, Betty Li, Germaine Fuh, Andrew Braisted, Marcus Ballinger, Ken Pearce, Manuel Baca, Lei Jin, Michelle Arkin, Greg Weiss, and Dan Erlanson. And thanks to everyone else in the Protein Engineering and Bioorganic Chemistry departments at Genentech: Bart de Vos, Henry Lowman, Tony Kossiakoff, Mark Ultsch, Bob Kelley, Nancy Gerber, Mike Randal, Melissa Starovasnik, Gerry Nakamura, and everyone else that helped me or just talked science with me in the hall, at beer hour, or at offsites. Thanks for support from the oligo synthesis, media prep, amino acid analysis, assay services, graphics, and fermentation groups at Genentech. Thanks Julie Ransom and Arlene Stuart, den mothers and friends. Thanks to a lively group of friends in the Biophysics department, especially Sarah Gillmor. Thanks to my collaborators, Paul Carter and John Ridgway. Thanks to my rotation advisors for teaching me a little something outside of protein engineering: Charley Craik, Tack Kuntz and Tom Scanlan. Thanks to the NIH Biotechnology training grant for supporting my first couple years, Genentech for support through most of my thesis work, and Sunesis Pharmaceuticals for support while I wrote.

Thanks to my thesis committee, Tom Scanlan, Wendell Lim, Bob Stroud and Jim Wells.

Thanks to Tack Kuntz for outlining the world of protein engineering during interviews and thereby convincing me that I should come to UCSF. Thanks also to Hong Xue and the scientists at Harvard who helped me traverse the distance from philosophy: Greg Verdine, Mike Eck, Steve Harrison and Don Wiley.

Thanks Mom, Dad and Janna.

Thanks to Damon Cole, Sarah Farivar, Pam Reinagel and Greg Bruell, Bill and Stephanie Dixon, and Micah Siegel, my close friends. Each of you know what you have meant to me. Thanks D. M. for showing me what is possible.

Those who dream by night in the dusty recesses of their minds wake in the day to find that it was vanity: but the dreamers of the day are dangerous men, for they may act their dream with open eyes, to make it possible.

T. E. Lawrence

Abstract

Phage display is a powerful technique for engineering proteins. Here we report the use of phage display to engineer protein-protein interfaces, catalysis and display itself.

The binding of human growth hormone (hGH) to its receptor is dominated by a small subset of the contact residues. The mutation W104A in the receptor reduces its affinity >2500-fold for hGH. Phage display was used to select a mutant of hGH that binds to the W104A mutant receptor. Three hGH mutations are critical for recovering binding to the W104A receptor: D171T, K172Y and E174A. The crystal structure of the mutant complex shows surprising local and global changes, including the rotation of several sidechains near the site of mutation and an overall rotation of the hormone relative to the receptor, inducing contact changes throughout the interface.

A complementary approach was applied to the C_H3 homodimer interface of the IgG heavy chain, in order to create heterodimers. Instead of mutating a large residue to a small one, as in the hGH receptor case, a small residue was mutated to a large one: T366W. Phage display was then used to accommodate the increased bulk in the interface creating a stable heterodimer.

Enzymes are attractive targets for protein engineering. Harnessing the power of phage display for enzyme engineering is demonstrated using a product capture strategy. Subtiligase, a double mutant of subtilisin BPN', that catalyzes the ligation of peptides was displayed on phage with an extended N-terminus. Active enzymes were selected from a pool of mutants by requiring that the enzyme-phage attach a biotin-peptide to their own N-termini. Mutagenesis of the active site yielded consensus mutations consistent with the catalytic mechanism of subtiligase and mutations that increased the rate of ligation.

Phage display is sometimes hampered by the difficulty of displaying certain proteins. It has been impossible to study hGH receptor using phage due to its poor display. In order to improve the display of the hGH receptor on phage for functional studies, DNA

shuffling was used in combination with binding selection. Mutants that had 60-fold improved display were selected. Characterization of the mutants showed that the mutations were concentrated in unstructured regions of hGHbp.

Abbreviations

Ab, antibody; Abz, 2-aminobenzoyl; cfu, colony forming units; cv, column volume; DTT, 1,4-dithio-D,L-threitol; ELISA, enzyme-linked immunosorbent assay; glc, glycolate; H, heavy; IA, immunoadhesin; L, light; MES, 2-(*N*-morpholino)ethanesulfonic acid; PBS, phosphate-buffered saline; PEG, polyethylene glycol; pfu, plaque-forming units; PMSF, phenylmethylsulfonyl fluoride; $Y(NO_2)$, 3-nitro-L-tyrosine; amino acids are indicated by single or triple letter code; mutations are designated by the single-letter amino acid code for the wild-type residue followed by its sequence position and the mutant residue, N- and C-terminal peptide substituents are indicated by chemical names hyphenated before or after the single letter peptide sequence.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Proteins are ubiquitous to life. They copy DNA, build cell walls, pass signals, manage energy, and make proteins. Therefore, protein engineering, the knowledge-based modification of proteins for particular goals, is tremendously valuable. The goal of protein engineering is to understand protein structure-function well enough to be able to design proteins with desired functions. The modularity of domains and homology between proteins makes swapping domains and fragments straightforward. However, modifying the binding specificity or catalytic activity of domains has been more challenging. Most protein design is "reasonable" rather than "rational". Experimental and theoretical information provides some predictive value. However, due to limitations in our understanding of these complex polymers, the effect of the mutations is very uncertain until actually tested. This uncertainty has promoted widespread use of random or semi-random mutagenic and selection techniques. The desired mutants can be obtained much faster if mutations are tested millions at a time rather than one at a time.

Phage display is the method of choice for identifying mutants that bind a target. The technique consists of fusing the protein of interest to a phage coat protein. The assembled phage then display the protein of interest on their surface. Mutants of the protein are selected by binding to a target and washing away the unbound phage. The phage displaying a tight binding mutant can then be multiplied by reinfecting bacteria and used for further rounds of binding selection and characterization.

In the following chapters I describe the use of phage display to study and engineer the interface between human growth hormone (hGH) and its receptor (hGHbp), the homodimer interface of the third constant domain of the antibody heavy chain (C_H3), the active site of subtiligase and the surface of hGHbp. In each case there was a specific engineering goal. At the same time, the phage selections were designed to shed light on functional relationships.

Several principles were highlighted. First, protein interfaces have considerable flexibility in the particular contacts that are made (Chapters 2 and 3). Second, deleterious mutations in protein interfaces can be compensated for, although sometimes in an unpredictable manner (Chapters 2 and 3). Third, phage display can be used to study and engineer enzyme activity (Chapter 4). Fourth, phage selection itself often gives clues to the effects of mutations on other protein properties, including folding and stability (Chapters 4 and 5).

Chapter 2. Structural Plasticity at a Protein-Protein Interface*

INTRODUCTION

Protein-protein interfaces are usually large and elaborate, consisting of 10 to 40 contact side chains, each of which interdigitates with several others across the interface (Janin & Chothia, 1990; Jones & Thornton, 1996). The contact side chains are often presented from discontinuous segments of each polypeptide chain. Given the complexities of these interactions we wondered how a functionally disruptive mutation on one side of the interface could be complemented by mutations in its binding partner. This is a challenge that nature faces as protein complexes co-evolve.

This problem was studied by phage display using the high-affinity interface between human growth hormone (hGH) and the extracellular domain of its receptor (hGHbp), a member of the cytokine receptor superfamily (Wells & de Vos, 1996). The hormone and receptor initially form a tight 1:1 complex (dissociation constant K_d = 0.3 nM), and the x-ray structure of this complex is known (Clackson *et al.*, 1998). There are about 30 contact side chains on each side of this interface, but alanine-scanning mutagenesis has shown that only a small set of primarily hydrophobic contacts at the center of the interface are responsible for the binding affinity (Cunningham & Wells, 1989; Cunningham & Wells, 1993; Bass *et al.*, 1991; Clackson & Wells, 1995). This energetic "hot spot" on the receptor includes a tryptophan residue (Trp104) that inserts into a groove in hGH (Figure 2-1A). When Trp104 is mutated to alanine, the binding affinity is >2500 times lower (Bass *et al.*, 1991; Clackson & Wells, 1995). We isolated a variant of hGH that could bind to W104A-hGHbp, and determined the x-ray structure of the complex to see how the interface adapted.

* Portions of this chapter appeared in Atwell, S., Ultsch, M., De Vos, A. M. and Wells, J. A. (1997) Structural plasticity in a remodeled protein-protein interface. *Science*, **278**, 1125-1128.

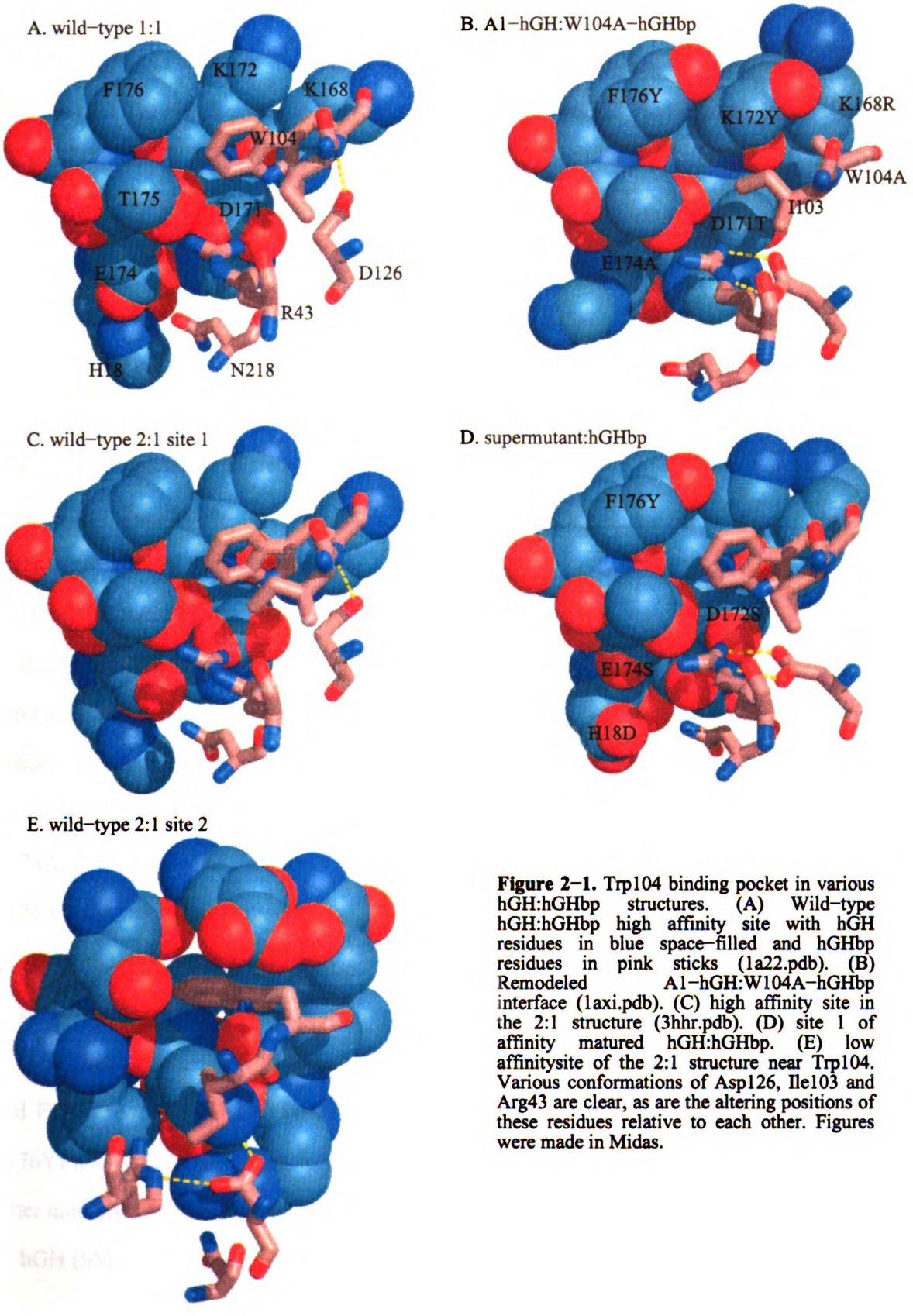

Figure 2–1. Trp104 binding pocket in various hGH:hGHbp structures. (A) Wild–type hGH:hGHbp high affinity site with hGH residues in blue space–filled and hGHbp residues in pink sticks (1a22.pdb). (B) Remodeled A1–hGH:W104A–hGHbp interface (1axi.pdb). (C) high affinity site in the 2:1 structure (3hhr.pdb). (D) site 1 of affinity matured hGH:hGHbp. (E) low affinitysite of the 2:1 structure near Trp104. Various conformations of Asp126, Ile103 and Arg43 are clear, as are the altering positions of these residues relative to each other. Figures were made in Midas.

RESULTS AND DISCUSSION

To isolate a mutant of hGH that could bind to the W104A-hGHbp, we randomly mutated hGH at the five residues (Lys168, Asp171, Lys172, Thr175, and Phe176) that pack against Trp104 (Figure 2-1A). A pool of hGH mutants, containing ~10^7 independent sequences, was displayed on filamentous phagemid particles and selected for binding to immobilized W104A-hGHbp. After seven rounds of binding selection, clones were sequenced and repeat isolates screened for binding to the W104A-hGHbp (Table 2-1). The most frequently isolated library member, called A1-hGH, was the only one found to bind detectably to the W104A-hGHbp in a screening assay. This variant contained five mutations (K168R, D171T, K172Y, E174A, F176Y). One of these, E174A, was not programmed by the mutagenic oligonucleotide. Such spurious mutations usually have a strong selective advantage (Cunningham *et al.*, 1994; Li *et al.*, 1995). Three of the selected residues (D171T, E174A and F176Y) were similar to previously selected mutations of hGH (D171S, E174S, and F176Y) that increased binding to the wild-type hGHbp (Lowman & Wells, 1993). These latter mutations had been combined with twelve other mutations to create a "supermutant" of hGH (SM) with dramatically increased affinity for hGHbp.

Table 2-1. hGH selectants for W104A receptor.

	randomized residue					BIAcore k_{off}[a]
clone	K168	D171	K172	T175	F176	
A1[b](4/1[c])	R	T	Y	.	Y	7×10^{-4}
A2	W	A	E	.	Q	ndb[d]
A3	R	G	V	E	E	
A4	G	E	W	A	E	
A5	Q	S	R	A	N	
A6	E	V	L	V	G	
A7	T	S	R	.	Y	
A8	Q	K	E	.	S	
A9	.	N	E	.	S	
A10[e](9/1)	.	.	.	.	.	ndb
A11[f](5/1)	R	.	.	.	.	ndb[g]

[a]off rate of protein from an *E. coli* shockate measured on BIAcore with random lysine coupled W104A hGHbp.
[b]this clone had a spurious E174A mutation.
[c]indicates number of observed clones and number of different DNA sequences.
[d]no detectible binding
[e]wild-type template contaminant.
[f]wild-type template with K168R.
[g]there appeared to be some binding of this mutant, but the off-rate could not be measured accurately.

The affinity for binding of the A1-hGH variant to the W104A-hGHbp was dramatically improved by the five mutations (K_d = 14nM) as compared to wild-type hGH binding to W104A-hGHbp (K_d >1000 nM) (Table 2-2). To identify the mutations in A1-hGH that most affected binding, we reverted each to their wild-type counterpart and measured their binding affinities to the W104A-hGHbp (Table 2-2). The K168R and F176Y mutations made only minor contributions to affinity: their reversion back to lysine and phenylalanine, respectively, only reduced affinity up to two-fold from the A1 variant (K_d = 33nM and 21nM, respectively). However, each of the other three selected mutations (D171T, K172Y and E174A) caused 8 to >300-fold reductions in affinity when reverted to their wild-type residue. A1-hGH bound much more weakly to the wild-type hGHbp showing it to be a specific binder of W104A-hGHbp. The reduced binding to wild-type hGHbp was due to Tyr172, since reverting to the wild-type Lys residue restored the wild-type affinity. This revertant also showed that the combined effect of the other four mutations increased affinity specifically for W104A-hGHbp, not for wild-type hGHbp. This is despite the fact that these mutations individually are homologous or identical to mutations known to increase the affinity of hGH for hGHbp.

Table 2-2. Reversion mutation of selectant A1.

hGH mutant	BIAcore vs. wt hGHbp[a]			RIA[b]	BIAcore vs. W104A hGHbp			RIA
	k_{on}[c]	k_{off}[d]	K_D[e] (nM)	K_D (nM)	k_{on}	k_{off}	K_D (nM)	K_D (nM)
wt	1.4×10^{5}[f]	3.7×10^{-4}[f]	2.6	0.31[g]	ndb[h]	ndb	ndb	>5000
knob	4.8×10^{5}	5.0×10^{-2}	100	>1000	2.0×10^{5}	2.2×10^{-3}	11	14
R168K	3.1×10^{5}	5.3×10^{-2}	170	300	1.7×10^{5}	4.8×10^{-3}	28	33
R168A	5.4×10^{5}	7.8×10^{-2}	140	200	2.0×10^{5}	3.7×10^{-3}	19	44
T171D	3.1×10^{5}	3.1×10^{-2}	100	75	1.8×10^{5}	7.0×10^{-2}	390	>5000
T171A	3.2×10^{5}	4.7×10^{-2}	150	200	2.7×10^{5}	7.3×10^{-3}	27	26
Y172K	3.2×10^{5}	7.6×10^{-4}	2.4	2.6	1.7×10^{5}	1.5×10^{-2}	88	200
Y172A	2.1×10^{5}	5.4×10^{-3}	26	[i]	1.4×10^{5}	1.5×10^{-2}	100	
Y172F	2.4×10^{5}	2.5×10^{-2}	100		1.6×10^{5}	8.8×10^{-4}	5.5	
A174E	ndb	ndb	ndb	>1000	1.3×10^{5}	3.7×10^{-2}	280	400
Y176F	2.9×10^{5}	2.2×10^{-2}	76	63	1.8×10^{5}	4.0×10^{-3}	22	21
Y176A	7.1×10^{4}[j]	1.6×10^{-2}	230		1.9×10^{5}	1.9×10^{-2}	100	
wt^{K168R}				0.7				>1000
SM^{k}				0.0009				1.6

[a]BIAcore measurement of on and off rates at 25 °C using hGHbp immobilized by random lysine coupling.

[b]radio immunoprecipitation assay using I^{125}-hGH. Errors were approximately 30% for the high affinity variants and as high as 100% for the affinities near the 1 μM upper limit of the assay.
[c]on rates determined by fitting line to apparent on rates at several hGH concentrations.
[d]off rates determine from dissociation of hGH variants equilibrated at approximately 1μM
[e]dissociation constant = k_{off}/k_{on}
[f]data for G120R-hGH binding to S201C-hGHbp from Cunningham & Wells (1993).
[g]from Cunningham & Wells (1989)
[h]no detectable binding
[i]not measured
[j]on rate for this mutant was determine with a single curve using BIAcore analysis software.
[k]affinity matured "supermutant" hGH (Lowman and Wells, 1993)

To understand the structural basis for these functional effects, we determined the x-ray structure of the complex between the A1-hGH and the W104A-hGHbp. The crystals were isomorphous with the wild-type 1:1 complex (Clackson *et al.*, 1998) and gave complete data to 2.1 Å, the highest resolution yet obtained for any hGH:hGHbp complex (Table 2-3). The structure shows that the 150 $Å^3$ hole in the W104A-hGHbp is filled by the phenolic group of the critical K172Y mutation of the A1-hGH (Figure 2-1B); in addition the Ile103 side chain in the W104A-hGHbp rotates around to place its isobutyl side chain in the cavity.

Table 2-3. Crystallographic statistics for the complex between the A1-hGH and the W104A-hGHbp.

Data	
Source	CHESS, beam line F1
Space group	$P4_32_12$
Unit cell (Å)	a=b=66.5 c=231.7
Resolution (Å)	100 to 2.1
R_{merge} (%)	5.7
Completeness (%)	91.8
Total reflections	133,933
Unique reflections	28,276
Redundancy	4.7
Anisotropic correction ($Å^2$)	12
Model refinement	
Resolution (Å)	7 to 2.1
Unique reflections	24,290
R.m.s.d. in bonds (Å)	0.014
R.m.s.d. in angles (°)	2.0
Average B-factor for protein atoms ($Å^2$)	27.7
R.m.s.d. in B-factor of bonded atoms ($Å^2$)	2.6
R (%)	19.0
R_{free} (%)	26.2
Residues	366
Waters	382

Asp126 breaks a hydrogen bond to the main chain at position 104 to form a salt bridge with Arg43 (Figure 2-1B). The movement of Asp126 is probably coupled to the critical D171T change in the A1-hGH, because retention of aspartate at position 171 would cause electrostatic repulsion at Asp126 of the W104A-hGHbp (Figure 2-1). Indeed, the binding of a T171A mutation is only two-times lower than that of A1-hGH, consistent with the

preference for a neutral residue over Asp at position 171 for binding to the W104A-hGHbp. The structure of an affinity matured hGH in complex with hGHbp has an analogous D171S mutation and a similar movement of Asp126 (Figure 2-1D). Two other structures show the additional variability in the positions of these receptor residues depending on what complex they are in: the wild-type hGH:hGHbp structure with two receptors bound per hGH (2:1) at the high affinity site (Figure 2-1C) has Ile103 rotated similarly to W104A-hGHbp, and the low affinity site has dramatically shifted receptor residues to accommodate binding to the back side of hGH (Figure 2-1E).

The E174A mutation in hGH induces another set of concerted changes (Fig 2-1B). In the wild-type complex, Glu174 makes two hydrogen bonds: one to His18 of hGH and the other to Asn218 of the hGHbp. When Glu174 is mutated to Ala, His18 rotates to avoid clashing with Asn218 which allows the hormone and receptor to move closer together (Figure 2-1B). The other two selected residues (K168R and F176Y) do not make new contacts with the hGHbp, consistent with their having little effect on binding when reverted back to their wild-type residues.

Overall, the combination of mutations with conformational changes nearly fill the 150 $Å^3$ cavity introduced by the W104A mutation. A hole of about 20 $Å^3$ hole is left between Tyr172 and Tyr176 of the hormone and Ile103 of the hGHbp; a hole of about 40$Å^3$ hole is created by the movement of the K168R side chain of hGH. These small cavities, which are not seen filled with ordered solvent, may be disruptive to affinity as was reported for mutations in T4 lysozyme (Ericksson *et al.*, 1992; Ericksson *et al.*, 1993).

Remarkably, the local structural changes near the site of the W104A mutation induce substantial global changes that affect the entire interface. The five mutations allow the hormone to move closer to the hGHbp in the vicinity of the W104A mutation, with the Cα of Lys168 1.1 Å closer to the Cα of W104A. This movement is accompanied by a 5° rotation of A1-hGH relative to the receptor. A1-hGH slides across the receptor pivoting at a point near Trp169 of hGHbp. The global changes that radiate outward from the pivot point

necessitate substantial rearrangement in both main-chain and side-chain contacts throughout the interface. Some of the main-chain and side-chain van der Waals and hydrogen bonding groups 15 Å distant from the epicenter of these mutations move by up to 3 Å relative to their neighbors in the wild-type hGHbp (Figure 2-2A). As a result, four hydrogen bonds are lost and three new ones are gained between unchanged atoms in the mutant complex. For example, the 3 Å salt bridge between Lys41 and Glu127 in the wild-type complex, which is 11 Å from the α-carbon of Trp104, has moved apart by 2.4 Å in the mutant complex. This gap is filled by two ordered waters, one of which mediates interactions between this charged pair. The salt bridge in the wild-type complex between Arg64 and Asp164, 17 Å from Trp104, is replaced by one between Arg64 and Glu44 in the mutant complex with commensurate changes in ordered waters.

By contrast, large domain movements and detailed changes in the interface were not seen in another mutant complex which included three alanine mutations in hGH known to have no effect on binding (F25A, Y42A, and Q46A) (Pearce *et al.*, 1996). These mutations, which are near the periphery of the interface, induced only small movements near the sites of mutation (Figure 2-2B); two hydrogen bonds were lost and none were gained.

The structures and binding regions for a number of hematopoietic cytokines and their receptors are now known (Wells & de Vos, 1996; Somers *et al.*, 1994; Walter *et al.*, 1995; Livnah *et al.*, 1996). One of the notable features is that although the same general regions of the hormone and receptor are involved in binding, the exact nature of the contacts is very different. For example, W104 of the hGHbp is conserved in only a few of the cytokine receptors, particularly the prolactin receptor, which shares 30% sequence identity with the hGHbp and also binds hGH. Interestingly, the structure of the complex between hGH and the prolactin receptor also shows substantial differences in domain orientation compared with hGHbp (Somers *et al.*, 1994).

These studies reveal how large functional changes can be rescued by a limited number of mutations. The mutations were accommodated by remarkably large movements in

unaltered side chain and main chain contacts throughout the interface. Structural changes induced by mutations have been seen in x-ray structures of single-domain proteins, for example when mutating core residues (Richards & Lim, 1994) or when inserting or deleting surface residues in α-helices of T4 lysozyme (Vetter *et al.*, 1996) or staphyloccal nuclease (Keefe *et al.*, 1993). Here, it is striking that such large changes are observed over a large interface in a protein complex, requiring coordinated changes in both components, and that they involve very specific interactions such as hydrogen bonds and salt-bridges. These results suggest that structural plasticity is nascent in protein-protein interfaces and provides a mechanism for mutations to be accommodated during the co-evolution of high affinity binding partners.

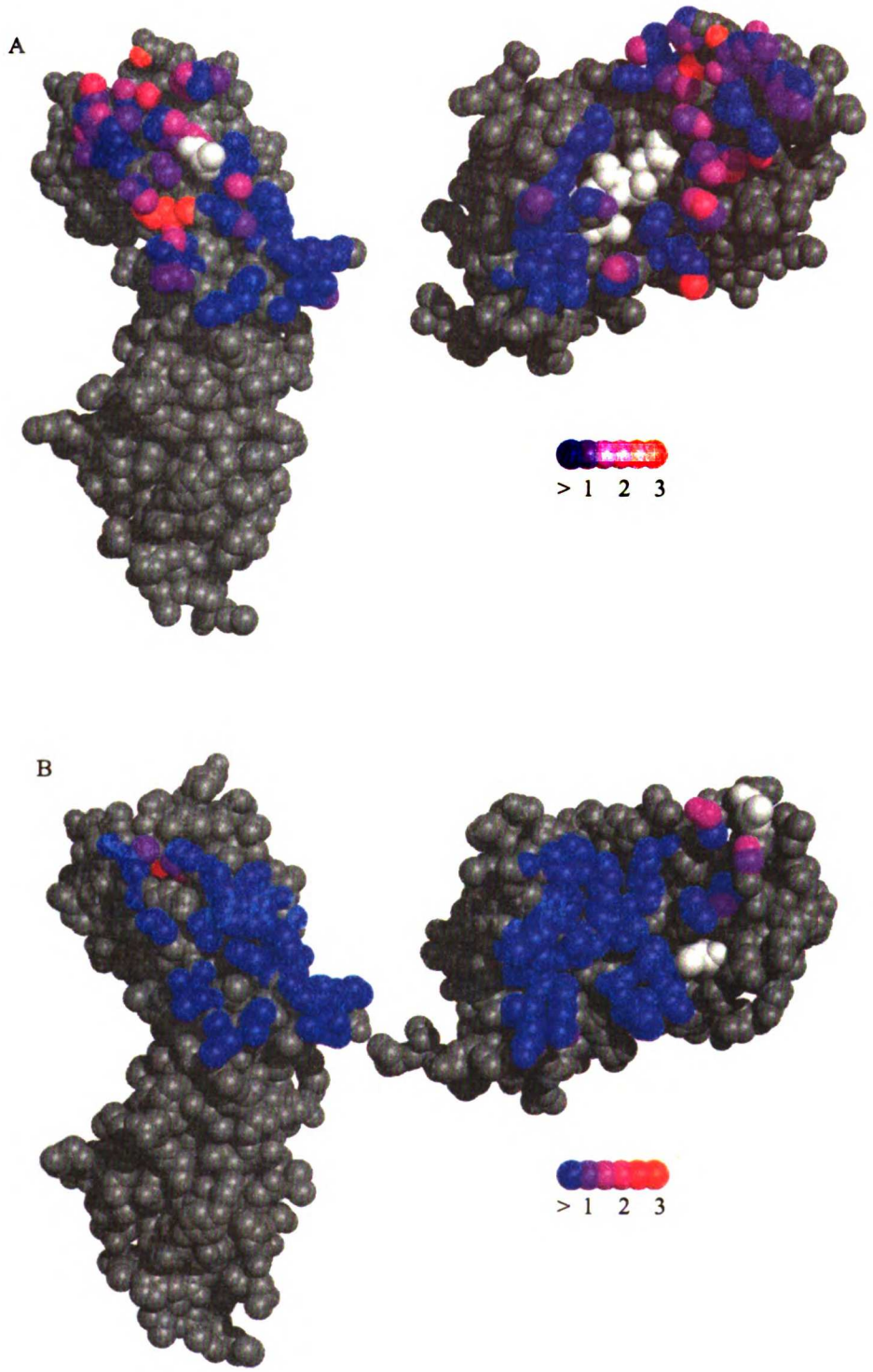

Figure 2–2. Movement of interface atoms in (A) the A1–hGH:W104A–hGHbp structure (1axi.pdb) and (B) the affinity inert mutant hGH:hGHbp (Pearce *et al.*, 1999). Atoms are colored on a sliding scale according to the average movement of their cross interface neighbor atom within 5 A. Non–contact residues are colored grey. Mutated residues are colored white and have truncated sidechains. Calculations were performed in the program Rigimol (DeLano, in preparation).

MATERIALS AND METHODS

Library. The library was constructed by site-directed mutagenesis (Kunkel *et al.*, 1991) by substituting NNS (N = A, G, C, or T; S = G or C) for the codon to be randomized. The library that randomized hGH at codons 168, 171, 172, 175, 176 was made by mutagenizing the plasmid (pH0753) with the oligonucleotide 5'-CTGCTCTACT-GCTTCAGGNNSGACATGNNSNNSGTCGAGNNSNNSCTTCGAATCGTGCAGTG-CCGCTCT-3'. The initial library size was $2x10^7$. The initial library had a bias toward wild-type template nucleotides, especially at nucleotides adjacent to non-mutagenized nucleotides (Table 2-4), probably due to the influence on annealing of nearby sequences. There was a large bias at codon 168. Nevertheless, mutants of the remaining residues should have been adequately sampled in the 10^7 library. Codon bias is much less of a problem when using stop templates. After characterizing mutant A1, we remade the library with a stop template but sorting yielded no consensus sequence even at residue 175 and many deletion mutants, a hallmark of unsuccessful selection.

Table 2-4. Nucleotide distribution from the initial library.[a]

	Codon														
	168			171			172			175			176		
A	**80**[b]	**75**	0	15	**30**	5	**30**	**35**	0	**60**	10	**20**	25	30	0
C	5	0	0	5	35	**35**	15	15	5	20	**40**	20	5	20	**60**
G	15	25	**100**	**75**	30	60	45	25	**95**	20	35	55	25	20	35
T	0	0	0	5	5	0	10	25	0	0	15	5	**45**	**30**	0

[a]percentage of particular base from 20 sequenced clones. Three contained partial wild-type and partial oligonucleotide derived sequence, one had a single nucleotide deletion.
[b]wild-type template nucleotides are in bold.

Libraries were also constructed for binding to W169A-hGHbp, which randomized residues 64 and 67 and 64, 67 and 68. After sorting seven rounds, libraries were dominated by stop template clones with no consensus among other clones.

The plasmid pH0753 was derived from phGHam-g3 (Lowman *et al.*, 1991) by mutation of the base at position 419 from C to G in order to eliminate an Xba I restriction site.

Phage. Phage displaying mutants of hGH were selected for binding to the W104A-hGHbp immobilized on an immunosorp plate (at 50 μg/mL, NUNC) by incubating overnight at 4° C in phosphate buffered saline solution (PBS (pH 7.4)). No significant enrichment versus an uncoated well was observed at any round. Sorting against the W104A-hGHbp was continued seven rounds before individual phagemids were sequenced.

Protein. The A1-hGH and single site revertants were constructed by site-directed mutagenesis of pB0720, a G120R-hGH expression plasmid. The mutants were expressed as soluble proteins (Chang *et al.*, 1987) in the strain 34B8, a non-suppresser strain of *E. coli*. Each mutant was isolated from 50 mL of culture in 250 mL shake flasks by pelleting the cells and freezing them at -20° C overnight. Cells were thawed and osmotically shocked in 10 mM Tris (pH 8.0). The concentration of hormone was determined by densitometry of Coomassie-stained gels (Cunningham & Wells, 1989; *ibid.* 1993) for the protein used in the RIA assays and by sandwich ELISA (by Genentech assay services) for the protein used in the BIAcore measurements. Both methods have errors of ±10%. Screening of selectants for binding was done by measuring the off rate on a BIAcore instrument by coupling the W104A-hGHbp through lysine residues to an RU of 2500 (Cunningham & Wells, 1989; *ibid.*, 1993; Johnsson *et al.*, 1991). RIA was performed for binding of hGH variants to the hGHbp or W104A-hGHbp as described (Spencer *et al.*, 1988). In the case of the wild-type hGHbp, labeled ^{125}I-hGH was used as competitor; for the W104A-hGHbp, an affinity optimized variant of hGH ("supermutant", SM) (Lowman & Wells, 1993) labeled with ^{125}I was used. This variant bound with an affinity of 1.6 (±0.7) nM to the W104A-hGHbp. Standard errors in the K_d measurements ranged from ±30 to ±100% of the value indicated in the text and were higher near the detection limits of the assay (~1μM).

Measurements by BIAcore of the on and off rates utilized W104A-hGHbp and hGHbp coupled by lysines to an RU of 2000 in separate channels of the same sensor chip. Sensograms were background subtracted using a sensogram from an unmodified channel. Off rates were determined from mutant sensograms at approximately 1 μM mutant using

the first 10% of dissociation. On rates were determined by fitting to the apparent on rate for multiple concentrations of hGH mutant.

Crystallography. A1-hGH and W104A-hGHbp were purified from *E. coli* fermentation culture using ion exchange and hydrophobic interaction columns. The two proteins were combined and purified as a complex on a gel filtration column. Crystallization was induced by streak seeding with crystals of the F25A/Y42A/Q46A/G120R-hGH:hGHbp complex (Pearce *et al.*, 1996) at approximately 5 mg/ml complex in 50 mM Bis-Tris (pH 6.5) and 17-20% saturated ammonium sulfate. Crystals were frozen in liquid nitrogen for data collection. 2.1 Å data were collected at CHESS using and ADSC CCD detector positioned at 175 cm with the direct beam in the corner and data was reduced using DENZO and SCALEPACK (Otwinowski, 1993). Solution of the structure was done by molecular replacement using coordinates from the 2.6 Å resolution 1:1 complex structure (Clackson *et al.*, 1998). The initial refinement was done in X-plor (Brunger *et al.*, 1987) and completed using the maximum likelihood method in the Refmac package of the CCP4 suite (Murshudov *et al.*, 1997) with building in O (Jones *et al.*, 1991) using sigma weighted maps. No σ cutoff was used. The final model includes the ordered residues for the A1-hGH (residues 3-48, 51-129, 136-148, 155-191) in complex with the W104A-hGHbp (residues 32-52, 61-72, 79-236).

Chapter 3. Stable Heterodimers from a Homodimer*

INTRODUCTION

The construction of protein heterooligomers from homooligomers is a useful tool for studying specificity and stability in protein-protein interactions (Robey & Schachman, 1985; Carter *et al.*, 1986; Bedouelle & Winter, 1986; Wente & Schachman, 1987). In addition, engineered heterooligomers such as bispecific antibodies (Ab) hold significant promise for human therapy (Carter *et al.*, 1995). Unfortunately, the use of designed heterooligomers has been hindered by the lack of methods to prepare such molecules.

The earliest routes to constructing heterooligomers relied upon cumbersome strategies to purify them free of contaminating homooligomers (Robey & Schachman, 1985; Carter *et al.*, 1986). Preparation of heterodimers has been made easier by engineering purification handles into one or both parent homodimers. For example, tagging one parent domain with a polyarginine tail allowed recovery of a heterodimer by ion exchange chromatography (Deonarain *et al.*, 1992). Incorporation of a polyhistidine tail into one chain and an epitope tag into the second chain enabled specific detection of a heterodimer in the presence of homodimers (Elsevier *et al.*, 1996).

Random association of monomers gives a maximum possible heterodimer yield of 50% (with 25% yield of each homodimer). A more efficient approach is to redesign domain interfaces to promote heterodimer formation and prevent the assembly of homodimers. Indeed, installation of a salt-bridge across a dimer interface was partially successful to this end (Ward *et al.*, 1987). A more successful strategy was the use of sterically complementary "knobs-into-holes" mutations to redesign Ab heavy (H) chains for

* Portions of this chapter appeared in Atwell, S., Ridgway, J. B. B., Wells, J. A. and Carter, P. (1997) Stable heterodimers from remodeling the domain Interface of a homodimer using a phage display library. *J. Mol. Biol.*, **270**, 26-35.

heterodimerization (Ridgway *et al.*, 1996). In this approach, a knob mutant was first obtained by replacement of a small amino acid with a larger one in the C_H3 domain. This type of mutation, since it introduces steric bulk into an already tightly packed interface, should obstruct homodimerization. A hole was then designed to accommodate the knob by replacement of a large residue with a small one in a coexpressed C_H3 domain. These mutations promoted heterodimerization of the heavy chains (Ridgway *et al.*, 1996). Knobs-into-holes mutations have also proved useful in enhancing the formation of a bispecific antibody fragment known as a diabody (Zhu *et al.*, 1997). In addition, high affinity receptors for small molecules have been obtained by rational design of each partner using sterically complementary mutations (Belshaw *et al.*, 1995).

Screening libraries of mutants offers a more comprehensive approach than rational design for remodeling domain interfaces. Phage display technology (Smith, 1985; Parmley & Smith, 1988) was used here to search such libraries for combinations of C_H3 domain interface mutations that give rise to the most stable heterodimers. Selected C_H3 heterodimers were compared in stability to a previously designed knob-into-hole heterodimer (Ridgway *et al.*, 1996).

RESULTS

Phage display system for C_H3 heterodimer selection. A phage display system was developed for the selection of stable C_H3 heterodimers (Figure 3-1). The selection uses a T366W mutant of C_H3 (Ridgway *et al.*, 1996) fused to a peptide affinity tag that is coexpressed with a second C_H3 fused to M13 gene III protein. A mutant library was created in this second C_H3 by randomization of the residues opposing the T366W mutation. Phage that displayed stable C_H3 heterodimers were then captured using an monoclonal antibody specific for the affinity tag.

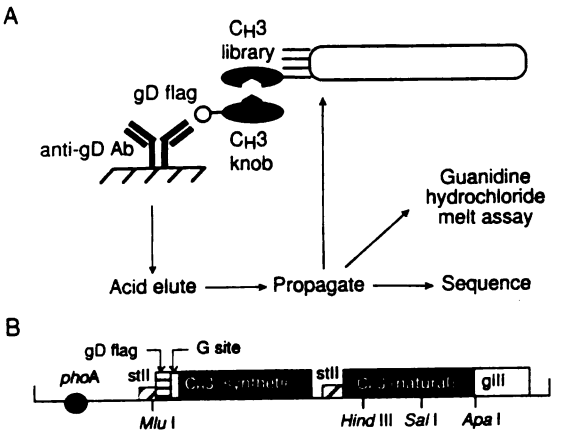

Figure 3-1. (A) Selection scheme for C_H3 heterodimer using phage display vector, pRA2. Phage displaying stable C_H3 heterodimers are captured using an antibody specific for the gD flag. (B) Schematic representation of a dicistronic operon in which C_H3 expressed from a synthetic gene is cosecreted with a second C_H3 expressed from the natural gene (Ellison *et al.* 1982) as a fusion protein with M13 gene III protein. Expression is under the transcriptional control of the alkaline phosphatase promoter (*pho*A, Chang *et al.*, 1986) and secretion is directed by the heat stable enterotoxin signal sequence (stII, Picken *et al.*, 1983). The synthetic C_H3 gene is preceded by a sequence encoding a peptide derived from herpes simplex virus glycoprotein D (gD flag, Lasky & Dowbenko, 1984; Berman *et al.*, 1985) and a cleavage site for the site-specific protease, Genenase I (Carter *et al.*, 1989).The knob mutation, T366W, is in the synthetic C_H3 gene, and the residues targeted for randomization are in the natural C_H3 gene (366', 368' and 407').

A C_H3 phage display library of 1.1 x 10^5 independent clones was constructed by replacement of a segment of the natural C_H3 gene with a PCR fragment. The fragment was obtained by PCR amplification using degenerate primers to randomize positions 366', 368' and 407'. The library was threefold larger than the number of codon permutations, implying a 96% theoretical coverage. The DNA sequences of 18 unselected clones were consistent with random codon usage and only one of these clones was a deletion mutant.

After two to five rounds of selection of the library the fraction of full-length clones was 90 %, 50 %, 50 % and 10 %, respectively, as judged by agarose gel electrophoresis of

single-stranded DNA with a concommitant increase in a shorter phagemid population. This problem of deletion mutants during selection was overcome by gel-purifying full-length phagemids and retransforming XL1-Blue cells.

Sequence analysis of deletion mutants. Eight out of 20 clones sequenced after round three were found to be deletion mutants that fused the gD flag directly to the C_H3-gIII fusion. Four mutants apparently arose by homologous recombination events between the C_H3 genes near the codons for residues Lys370, Tyr391, Leu398 and Tyr436. This was in spite of our efforts to minimize the risk of recombination by lowering the sequence identity between C_H3 genes to ~80% by using different codons in the synthetic and natural genes and by using a recombination-deficient (*rec*A) *E. coli* host, XL1-Blue. The other four deletion mutants had recombined between the end of the synthetic C_H3 gene and the end of the gly-ser linker of gIII at a homologous CCGG, thereby deleting the natural C_H3 gene and the gly-ser linker. It may be possible to reduce the frequency of homologous recombination between C_H3 genes by making their DNA sequences as divergent as possible while maintaining their protein sequence identity. The selective advantage of the deletion mutants might be minimized by reversing the order of gD-C_H3 and C_H3-gIII fusions. In this configuration intramoleculer loop-out deletions will remove gIII and the gD flag. This would prevent corresponding phage from surviving selective panning using the anti-flag Ab.

Deletion mutants might out-compete full-length clones in the propagation and/or panning steps. The relative importance of these two possible mechanisms remains to be determined. The 3,000-fold enrichment measured for the round five library and 27,000-fold for round seven, which were mostly deletion mutants, indicates that deletion mutants had a tremendous advantage in the binding selection.

Despite the problem of deletions arising during panning it seemed unlikely that stable heterodimers would have been completely lost from competition with the deletion mutants. A mean of $>10^6$ copies of each clone was used per round of panning. The sites of deletion

were scattered (see above), suggesting that they were not strongly influenced by the particular mutations. Thus, numerous copies of each clone in the library are anticipated to have been available for selection.

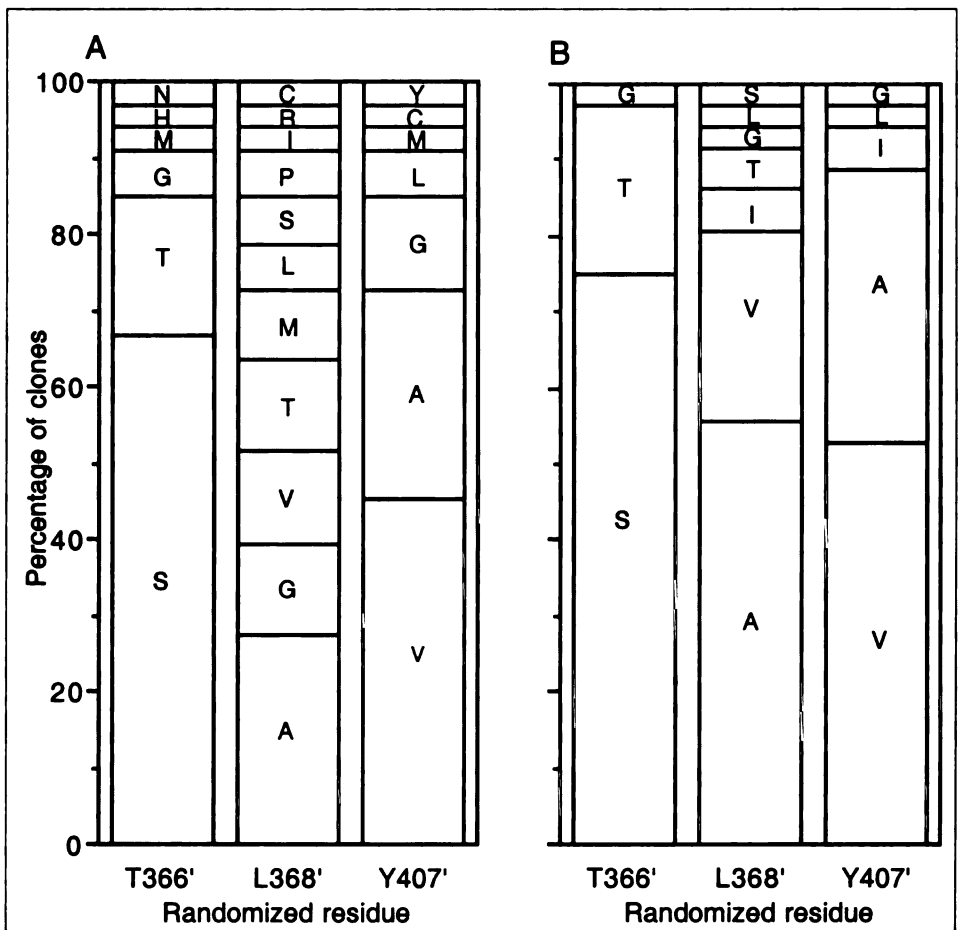

Figure 3-2. Frequency distribution of amino acids found at positions 366', 368' and 407' (A) following five rounds of panning from 33 clones sequenced and (B) seven rounds from 37 clones.

Analysis of selectants from C_H3 library. After purifying full-length phagemids and sequencing 33 from round five no single sequence appeared more than thrice, but some consensus was evident (figure 3-2A). Two additional rounds of selection were performed and the full length phagemids again purified and used to transform XL1-Blue cells before analyzing individual clones. The C_H3 mutants obtained after seven rounds of panning approached a consensus amino acid sequence at the randomized residues (Figure 3-2B).

Virtually all clones had serine or threonine at residue 366' indicating a very strong preference for a β-hydroxyl at this position. A strong preference for hydrophobic residues was observed for residues 368' and 407', with valine and alanine predominating. Six different amino acid combinations were recovered at least twice, including the triple mutant, T366'S/L368'A/Y407'V, which was recovered 11 times (Table 3-1). None of the sequenced selectants from any round was identical to a previously designed heterodimer, T366W:Y407′A (Ridgway *et al.*, 1996). The phage selectants may be less tightly packed than the wild-type C_H3 homodimer, as judged by a 40 to 80 $Å^3$ reduction in total side chain volume of the domain interface residues (Table 3-2).

Table 3-1. Comparison of C_H3 variants

	Residue identity				Selectant frequency	Different DNA sequences	Δ side chain volume ($Å^3$)[b]	[Guanidine]$_{1/2}$ (M)[c]
		randomized						
variant	366	366'	368'	407'				
wild-type	T	T	L	Y	na	na	(0)	> 3
designed[a]	W	T	L	A	na	na	− 4	< 0.5
selected	W	S	A	V	11	5	− 43	1.7
selected	W	S	V	A	4	3	− 43	1.6
selected	W	T	A	A	3	3	− 61	1.9
selected	W	S	A	A	3	3	− 81	1.5
selected	W	S	G	V	2	2	− 62	1.1

[a] From Ridgway *et al.* (1996).

[b] The change in total residue volume at the interface relative to wild-type C_H3 using volumes defined by the van der Waals radii (Creighton, 1993).

[c] The concentration of guanidine hydrochloride that resulted in a 50 % reduction in ELISA signal (see Materials and Methods). Identical ranking of clones was obtained in duplicate experiments, whereas [guanidine]$_{1/2}$ varied by ± 0.2 M.

na, not applicable.

Stability of C_H3 heterodimers on phage to guanidine hydrochloride. The stability of C_H3 heterodimers was first assessed by titrating corresponding phage with guanidine hydrochloride, incubating 5 minutes, followed by dilution and quantification of residual heterodimer by enzyme-linked immunosorbent assay (ELISA). The most frequently recovered heterodimer, T366W:T366′S/L368′A/Y407′V, is similar in stability to other

phage-selected heterodimers (Table 3-2). This phage-selected heterodimer is significantly more stable than the designed heterodimer, T366W:Y407'A but less stable than the wild-type C_H3 homodimer (Figure 3-3). The wild-type homodimer signal did not exhibit a cooperative unfolding transition in increasing guanidine concentrations. Interestingly, this same linear response to denaturant was observed for all five of the selectants in urea, indicating that these proteins respond differently to the denaturants.

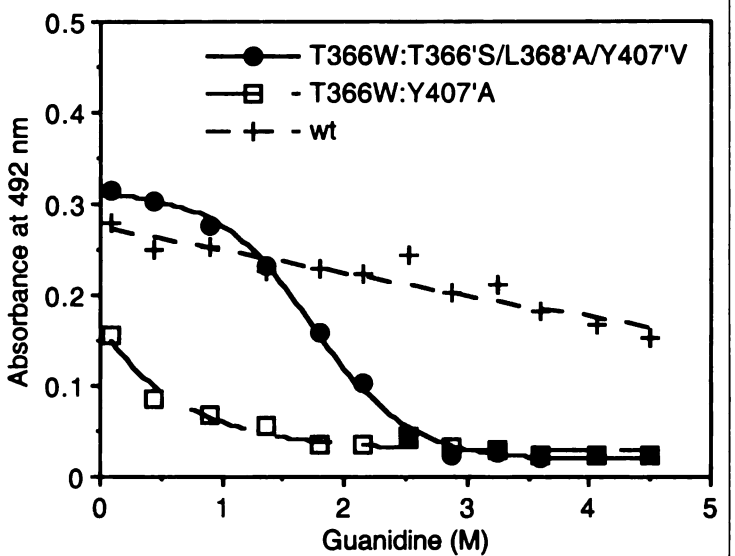

Figure 3-3. Guanidine hydrochloride denaturation of C_H3 heterodimers on phage. Phage were incubated with denaturant for 5 min, diluted 10-fold and the residual C_H3 heterodimer estimated by ELISA (see Materials and Methods).

C_H3 purification. C_H3 variants were secreted from *E. coli* grown to high cell density in a fermentor. The T366S/L368A/Y407V mutant purified by DEAE-Sepharose FF, ABx and Resource S chromatography gave a single major band following SDS-PAGE. Other C_H3 variants were recovered with similar purity. The C_H3 mutants were recovered in significantly lower yield (13 to 75 mg/l) than the for the wild-type C_H3 (310 mg/l). The molecular masses of wild-type C_H3 and T366S/L368A/Y407V, T366W and Y407A variants determined by high resolution electrospray mass spectrometry are within 1 au from those expected.

Differential scanning calorimetry. The guanidine hydrochloride denaturation assay with C_H3–phage provides a means to screen selectants rapidly. The most frequently recovered heterodimer, T366W:T366'S/L368'A/Y407'V, is also one of the most stable. This phage-selected heterodimer was chosen for more detailed characterization by

differential scanning calorimetry including comparisons with the designed heterodimer, T366W/Y407′A, and the wild-type C_H3 (Figure 3-4). All C_H3 variants, both individually and in combination, were found to be dimers by size exclusion chromatography under the conditions that these same molecules were studied by calorimetry (1.75 mg/ml, in phosphate-buffered saline (PBS)). The only exception was the T366'S/L368'A/Y407'V mutant alone which had a slightly shorter retention time than C_H3 dimers.

A 1:1 mixture of T366W knob and T366S/L368A/Y407V hole mutants melts with a single transition at 69.4 °C, consistent with subunit exchange and formation of a stable heterodimer. In contrast, the T366W knob homodimer is much less stable than the T366W:T366′S/L368′A/Y407′V knob-into-hole heterodimer (Δt_m −15.0 °C). The T366S/L368A/Y407V hole mutant on its own is prone to aggregate upon heating and does not undergo a smooth melting transition.

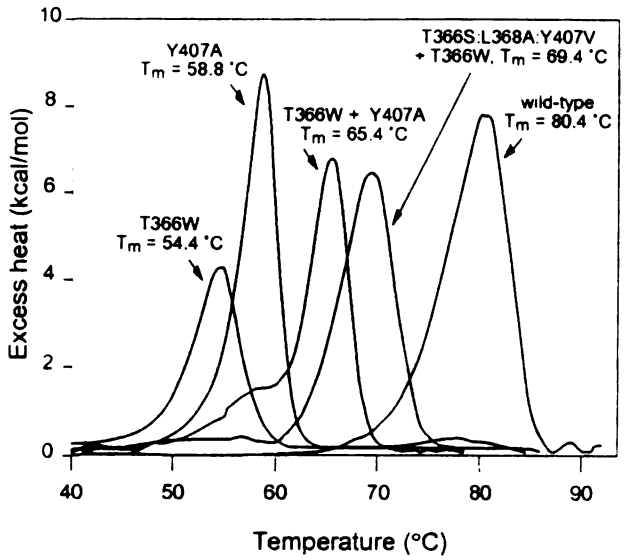

Figure 3-4. Thermal denaturation of C_H3 variants. The excess heat profiles observed upon heating 1.75 mg/ml (0.15 mM) solutions of C_H3 variants in PBS at 1°C/min.

The designed hole mutant, Y407A, melts at 58.8 °C and 65.4 °C in the absence and presence of the T366W knob mutant, respectively. This is consistent with subunit exchange and formation of a T366W:Y407′A heterodimer that has greater stability than either T366W (Δt_m = 11.0 °C) or Y407A (Δt_m = 6.6 °C) homodimers. The phage-selected heterodimer, T366W:T366′S/L368′A/Y407′V, is more stable than the designed heterodimer, T366W:Y407′A, (Δt_m = 4.0 °C), but is less stable than the wild-type C_H3 homodimer (Δt_m = −11.0 °C). Identical t_m values were found for first and second melts for all homodimers and

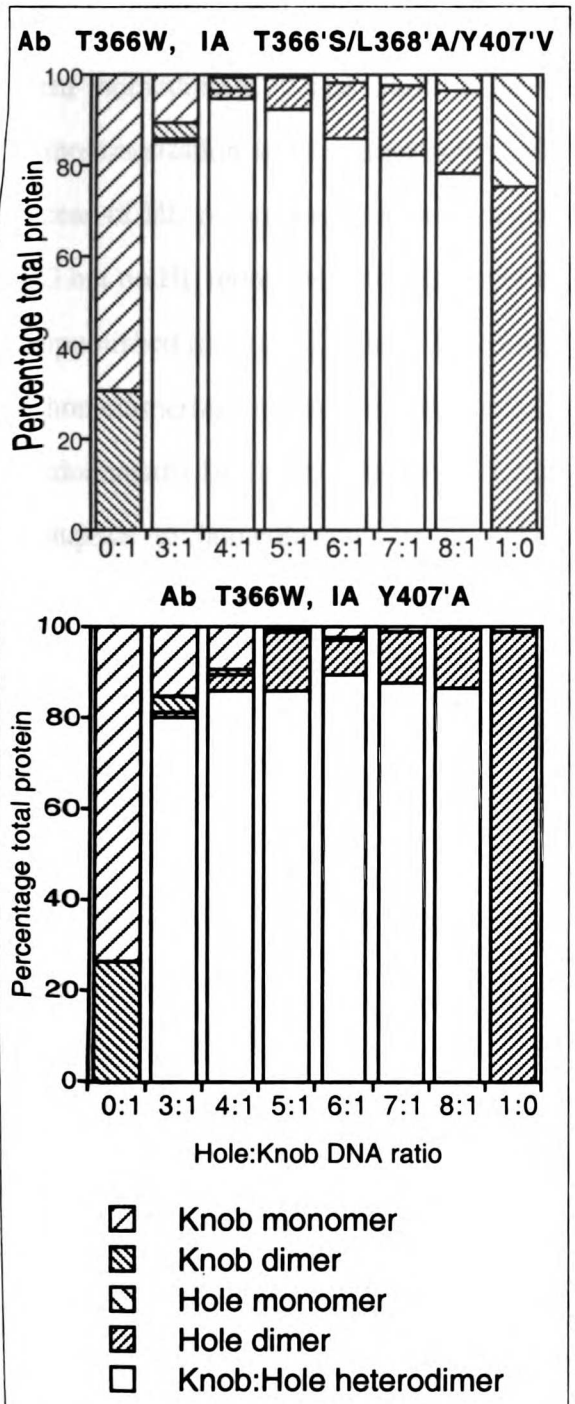

Figure 3-5. Scanning densitometric analysis of SDS-PAGE of protein A purified products from cotransfection of Ab H and L chains with IA. Data presented are the mean from 2 independent experiments.

heterodimers for two independent experiments (not shown). Thermal denaturation of C_H3 variants is not fully reversible as judged by diminished excess heat capacity upon cooling and reheating the sample and the appearance of turbidity.

Ab/IA cotransfection assay. Phage-selected and designed C_H3 mutants were compared in their ability to direct the formation of an Ab/IA hybrid, anti-CD3/CD4-IgG (Chamow *et al.*, 1994). This was accomplished by coexpression of humanized anti-CD3 light (L) and H chains together with CD4-IgG. Formation of heterodimers and homodimers was assessed by protein A purification followed by SDS-PAGE and scanning laser densitometry (Ridgway *et al.*, 1996). Comparable yields of Ab/IA hybrid were recovered from cotransfections in which the anti-CD3 H chain contained the designed knob mutation, T366W, and the IA contained either the phage-selected mutations, T366'S/L368'A/Y407'V, or designed hole mutation, Y407'A (Figure 3-5).

Phage-selected and designed C_H3 mutants were next evaluated in their propensity to form homodimers. The knob mutation, T366W, is apparently very disruptive to homodimerization since cotransfection of corresponding Ab H and L chains leads to an excess of HL monomers (may include non disulfide-bonded IgG) over IgG. In contrast, IgG but no HL monomers are observed for the same Ab containing wild-type C_H3 domains (unpublished results). The hole mutations, T366S/L368A/Y407V, are somewhat disruptive to homodimerization since transfection of the corresponding construct leads to a mixture of predominantly IA dimers with some IA monomers. The hole mutation, Y407A, is minimally disruptive to homodimerization as judged by the presence of IA dimers but no IA monomers following transfection of the corresponding phagemid.

DISCUSSION

Phage system allows selection of stable heterodimers. A phage display selection system was developed as a comprehensive approach to remodeling the domain interface of C_H3 homodimers. This system allows the selection for C_H3 mutants that form stable heterodimers and possibly selection against mutants that form stable homodimers. The counter selection against homodimers occurs because "free" C_H3 mutants will compete with the flagged C_H3 knob mutant for binding to available C_H3 mutant–gene III fusion protein. The free C_H3 mutants arise as a result of the amber mutation between the natural C_H3 gene and M13 gene III. In an amber suppressor host such as XL1–Blue, both C_H3-gene III fusion protein and corresponding free C_H3 will be secreted.

The selection relies upon the concentration of C_H3 being low during phage propagation, allowing dissociation of unstable heterodimers, plus plate binding and the dissociation of the least stable heterodimers. The concentration of phage in culture supernatants is $\leq$ 2 nM ($\leq 10^{12}$ cfu/ml) and the concentration of heterodimer is anticipated to be $\geq$ 10-fold lower from experience with other proteins with this monovalent display system (Bass *et al.*, 1990; Lowman *et al.*, 1991).

Guanidine hydrochloride denaturation proved to be a useful tool for the preliminary screening of the stability of C_H3 heterodimers on phage. It may be possible to use guanidine to increase the stringency of selection, since phage maintain infectivity for *E. coli* even after exposure with up to 5 M guanidine hydrochloride (Figini *et al.*, 1994).

Structural implications of C_H3 domain interface mutations. Residue T366 in the wild-type C_H3 domain forms a hydrogen bond with Tyr407' on the partner C_H3 domain (Miller *et al.*, 1990; Figure 7A). Maintaining a hydrogen bond at this position is apparently highly favored since serine or threonine was recovered at this position in virtually all of the phage selectants.

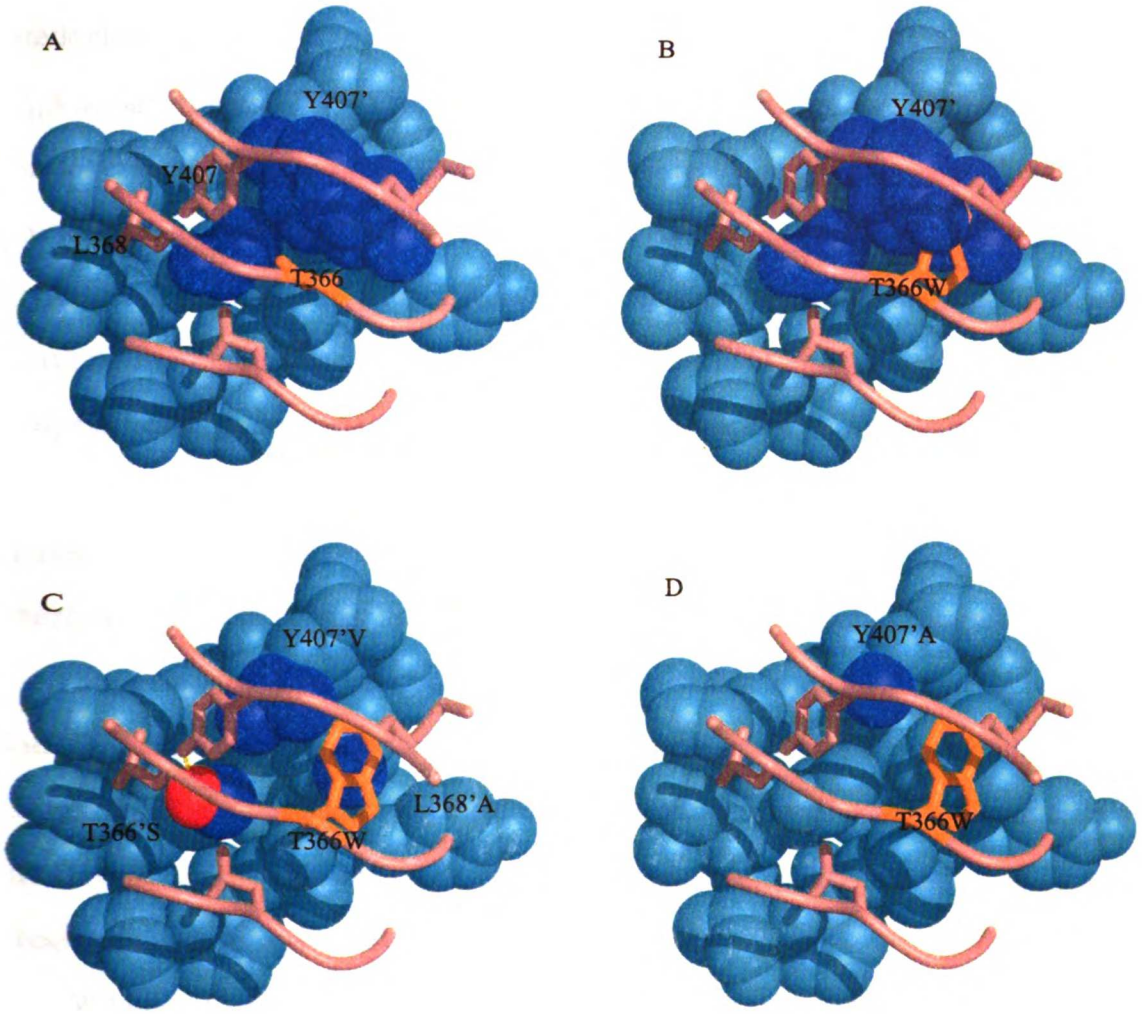

Figure 3–6. CH3 domain interface surounding residue 366. (A) 2.9 A structure of the wild–type CH3 homodimer (Deisenhofer, 1981). Residues on one side of the interface are in pink sticks and residues on the other side in blue spheres. Residues targeted for mutagenesis are in dark blue. (B) T366W modeled in the wild–type structure in the rotomeric position that clashes least with other residues on the same side of the interface. (C) Model of T366W:T366'S/L368'A/Y407'V showing the tryptophan fitting nicely in the hole left by the mutations. (D) Model of the designed T366W:Y407A showing a residual clash between Trp366 and Leu368'. Models were generated in Insight II using Biopolymer.

Computer graphics modeling of the C_H3 knob mutant, T366W, suggests that a potential steric clash may occur with residues T366′, and especially L368′ and Y407′ on the partner C_H3 domain (Figure 7B). This was the basis of our choice to target these residues for randomization. This putative steric clash is apparently relieved by the phage-selected mutations T366'S/L368'A/Y407'V (Figure 3-6C). Determination of the three dimensional structures of corresponding heterodimers will be required to elucidate the structural changes resulting from the domain interface mutations. Structural comparison of designed and phage-selected heterodimers may facilitate further rational design of domain interfaces.

Surprisingly, the T366W mutant is capable of forming homodimers, as judged by gel filtration, despite the anticipated steric clash. At least local rearrangement of the C_H3 domain interface must occur in order to accommodate the two extra tryptophan residues.

Phage-selected C_H3 heterodimers are more stable than designed heterodimer. The C_H3 heterodimers selected using phage are more stable than the designed heterodimer, T366W:Y407′A, to guanidine hydrochloride denaturation (Table 3-2). This difference in stability may account for why the designed C_H3 heterodimer was not amongst the sequences selected from the phage library.

Wild-type C_H3 domains exist exclusively as dimers at concentrations as low as 0.1 nM (Ellerson *et al.*, 1976), suggesting that the dissociation constant is < 0.1 nM. Our attempts to determine dissociation constants for different C_H3 homodimers and heterodimers by surface plasmon resonance were confounded by monomer-dimer equilibria in solution.

The phage-derived heterodimer, T366W:T366'S/L368'A/Y407'V, enhances the yield of Ab/IA hybrid in the cotransfection assay marginally compared to the designed heterodimer, T366W:Y407'A. This is in contrast to its increase in stability. We hypothesize that the yield of Ab/IA hybrid reflects primarily the relative rates of association of chains into homodimers and heterodimers rather than the thermodynamic stability of Ab/IA hybrids

(Ridgway *et al.*, 1996). This is a consequence of hinge disulfide bond formation, which restricts free exchange of partners.

Comparison of rational design and screening of libraries of mutants. Rational design and screening of phage display libraries are complementary approaches to remodeling a domain interface of a homodimer to promote heterodimerization. In the case of C_H3 domains, designed mutants identified domain interface residues that could be recruited to promote heterodimerization. Phage display was then used here to search permutations of three residues neighboring a fixed knob for combinations that most efficiently form heterodimers. Phage selectants might potentially be used to facilitate further rational redesign of the domain interface.

Phage selection strategies are useful for remodeling protein-protein interfaces. The phage remodeling of the hGH:hGHbp interface (Chapter 2) was very successful. A rational domain interface engineering strategy was recently used to design inactive mutants of HIV protease that would not readily homodimerize but would efficiently heterodimerize with the wild-type protease, thereby inactivating it (McPhee *et al.*, 1996). It may be possible to further optimize HIV protease heterodimers using a phage selection system similar to that developed here.

MATERIAL AND METHODS

Materials. Resource S and Superdex 75 HR 10/30 columns, DEAE-Sepharose-FF resin, and anti-M13 horseradish peroxidase conjugate were purchased from Pharmacia Biotech. Abx and ProSep A chromatography media were obtained from J. T. Baker and BioProcessing Ltd, respectively. Reagents for dideoxynucleotide sequencing were from United States Biochemicals. Centriprep-10 concentrators were purchased from Amicon. Polyacrylamide gels were obtained from Novex. The anti-gD Ab, 5B6 (Lasky & Dowbenko, 1984; Berman *et al.*, 1985) was generously provided by Brian Fendly (Genentech, Inc.). *E. coli* strain XL1–Blue was obtained from Stratagene. *o*-phenylenediamine was purchased from Sigma. Maxisorp immuno plates were purchased from Nunc.

Construction of C_H3 phage display vector. A synthetic gene encoding the C_H3 domain of human IgG_1 was assembled from 14 synthetic oligonucleotides 53 to 74 in length with 4 bp overlaps with neighboring oligonucleotides. Codons were chosen firstly to create unique restriction sites and to avoid self-complementary sequences predicted to form stable hairpins (≥ 14 Kcal/mol) using the program dyads (Martinez *et al.*, 1988). A secondary selection of codons was made to use those commonly found in *E. coli* proteins (Aota *et al.*, 1988) and to avoid codons found at corresponding positions in the natural C_H3 gene. The assembled gene was ligated with the large MluI/ApaI fragment from phagemid pmy95 (Garrard *et al.*, 1991). A correctly assembled gene was identified by dideoxynucleotide sequencing (Sanger *et al.* 1977) using Sequenase version 2.0. A phage display vector, pRA1, for cosecretion of a gD peptide flag-C_H3 fusion protein with a C_H3-M13 gene III fusion protein was assembled by site-directed mutagenesis (Kunkel *et al.*, 1987) of the synthetic C_H3 gene, and subcloning of a PCR-amplified natural C_H3 gene fragment (Ellison *et al.*, 1982) using standard methodology (Ausubel *et al.*, 1996).

C_H3 phage library construction and selection. Phagemid pRA2 for library construction was created by site-directed mutagenesis of pRA1 (Figure 3-1). Briefly, two *Hin*dIII sites and a SalI site were first removed, unique SalI and HindIII sites were then installed in the natural C_H3 gene, and a mutation encoding the T366W replacement made in the synthetic C_H3 gene. The small SalI/HindIII fragment from pRA2 was replaced with a PCR fragment generated by amplification of the natural C_H3 gene (Ellison *et al.*, 1982) with degenerate primers designed to randomize residues 366', 368' and 407': 5'-CCCTGCTG-CCACCTGCTCTTGTCGTCGACGGTGAGCTTGCTSNNGAGGAAGAAGGAGCCG-3' and 5'-GAGATGACCAAGAACCAGGTAAGCTTGNNSTGCNNSGTCAAAGGCT-TCTAT-3', where N = G, T, C or A and S = G or C.

Phage were prepared from the library in XL1-Blue by infection with M13KO7 helper phage (Vieira & Messing, 1987). Phage were then panned as described previously (Lowman, 1996) with the following modifications. Maxisorp immuno plates were coated overnight with 5 μg/ml 5B6 in 50 mM Na_2CO_3 (pH 9.6). Plates and phage were separately preincubated with 5 % (w/v) skimmed milk in phosphate-buffered saline (PBS) for 30 min at ~25 °C prior to coincubation for 2 h. The plates were washed 10 to 30 times with 0.1 % tween 20 in PBS, eluted with 200 mM glycine (pH 2.0) and then neutralized Tris Base. Fresh XL1–Blue cells were infected with the eluted phage and then superinfected with M13KO7 to prepare phage for the next round of selection. Enrichment was monitored by comparing the titer of phage eluted from 5B6-coated plates with that eluted from uncoated plates. The presence of deletion mutants after each round of selection was evaluated by agarose gel electrophoresis of phage in the presence of SDS and by dideoxynucleotide sequencing. After 5 rounds of selection full-length phagemids were separated from deletion mutants and used to retransform XL1–Blue. The transformants were subjected to 2 additional rounds of panning and full-length phagemids again used to transform XL1-Blue.

Individual clones after 5, 6 and 7 rounds of panning were analyzed by dideoxynucleotide sequencing.

C_H3-phage guanidine denaturation assay. Phage were prepared from individual clones following 7 rounds of selection and also from the parent vector, pRA1. Briefly, phagemids in XL1-Blue were used to inoculate 25 ml LB broth containing 50 μg/ml carbenicillin and 10 μg/ml tetracycline in the presence of 10^9 pfu/ml M13KO7 and incubated overnight at 37°C. The cells were pelleted by centrifugation (6000 *g*, 10 min, 4 °C). Phage were recovered from the supernatant by precipitation with 5 ml 20 % (w/v) PEG, 2.5 M NaCl followed by centrifugation (12000 *g*, 10 min, 4 °C) and then resuspended in 1 ml PBS. 180 μl 0–6 M guanidine hydrochloride in PBS was added to 20 μl phage preparations and incubated for 5.0 min at ~25 °C. Aliquots (20 μl) of each phage sample were then diluted 10-fold with water. The presence of C_H3 heterodimer was assayed by ELISA using 5B6-coated plates and detecting the phage with an anti-M13 polyclonal Ab conjugated to horseradish peroxidase, using *o*-phenylenediamine as the substrate. The reaction was quenched by the addition of 50 μl 2.5 M H_2SO_4 and the absorbance measured at 492 nm. The absorbance data were plotted against the guanidine hydrochloride concentration during the melt and fitted to $y=a+(b-a)/(1+10^{c(d-x)})$, (where y is the ELISA signal, a and b are maximum and minimum plateau values, x is the guanidine concentration and d is the $Gnd_{1/2}$) by a non-linear least squares method using Kaleidagraph 3.0.5 (Synergy Software).

Expression and purification of C_H3 variants. Selected C_H3 variants were subcloned into the expression plasmid pAK19 (Carter *et al.*, 1992) and transformed into *E. coli* strain, 33B6 (Rodrigues *et al.*, 1995), 33D3 (W3110 *ton*A *ptr*3 *pho*AΔE15 *lac*I^q *lac*L8 *deg*P *kan*R), or 17B8 (W3110 *ton*A *pho*AΔE15 Δ(*arg*F–*lac*)169 Δ(*srl*R–*rec*A)306). C_H3 variants were then secreted from corresponding *E. coli* grown for approximately 48 h at 30 °C in an aerated 10 liter fermentor as described previously (Carter *et al.*, 1992). C_H3

variants were purified by thawing 100-150 g fermentation paste in the presence of 300 ml 25 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1 mM PMSF, and 0.5 mg/ml lysozyme over ~30 min with vigorous mixing. Cell debris was removed by centrifugation (30000 *g*, 20 min, 4 °C). The resultant supernatant was dialyzed against 25 mM Tris-HCl (pH 7.5) and then loaded onto a 5 x 20 cm column of DEAE-Sepharose FF equilibrated with the same buffer. The resin was washed with 3 column volumes (cv) of 25 mM Tris-HCl (pH 7.5), prior to elution with a linear gradient of 0–100 mM NaCl in 25 mM Tris-HCl (pH 7.5), over 1.5 cv. Pooled fractions were adjusted to 5 mM EDTA and 0.1 mM PMSF, dialyzed against 25 mM MES (pH 5.5) and then loaded onto 2.5 x 15 cm column of ABx resin equilibrated with the same buffer. The column was washed with 5 cv 25 mM MES (pH 5.5), prior to elution with a linear gradient of 0–150 mM $(NH_4)_2SO_4$ in 25 mM MES (pH 5.5) over 10 cv. Pooled fractions were dialyzed against 25 mM CH_3COOK (pH 4.6) and then loaded on to a 6 ml Resource S column equilibrated with the same buffer. The column was washed with 5 cv 25 mM CH_3COOK (pH 4.6) and then eluted with a linear gradient of 0-350 mM NaCl in 25 mM CH_3COOK (pH 4.6) over 20 cv. The purified C_H3 was concentrated using Centriprep-10 concentrators, flash frozen in liquid nitrogen and stored at –70 °C. The yield of C_H3 variants was estimated from the A_{280} and the corresponding experimentally determined extinction coefficients. Individual C_H3 variants together with phage-selected and designed knob-into-hole mutant pairs (total concentration of 1.75 mg/ml) were analyzed by size exclusion FPLC with a Superdex 75 HR 10/30 column using PBS and a flow rate of 1.0 ml/min.

Differential scanning calorimetry. Thermal denaturation experiments were performed on a Microcal (Northampton, MA) MC-2 differential scanning calorimeter. C_H3 samples were prepared by extensive dialysis at ambient temperature against PBS. The T366W mutant was mixed separately with T366'S/L368'A/Y407'V and Y407A mutants at a molar ratio of 1:1 and at a total concentration of 1.75 mg/ml then incubated overnight at

ambient temperature prior to calorimety. Individual C_H3 variants and mixtures (1.75 mg/ml) were heated to 95 °C at 1 °C/min. The observed melting profiles were baseline–corrected and normalized using the software provided by Microcal (version 2.9). Data were separately fitted to various single or two transition models for unfolding.

Liquid chromatography–mass spectrometry. C_H3 variants in PBS were mass analyzed with a packed capillary liquid chromatography system coupled directly to a PE-Sciex API 3 triple quadrupole mass spectrometer (Sciex, Toronto, Canada) as described previously (Bourell *et al.*, 1994).

Expression and purification of Ab/IA variants. Phagemids encoding anti-CD3 H and L chains (Shalaby *et al.*, 1992; Rodrigues *et al.*, 1992) and a CD4-IgG IA (Byrn *et al.*, 1990) incorporating C_H3 mutations were cotransfected into human embryonic kidney cells, 293S, for transient expression as described (Chamow *et al.*, 1994; Ridgway *et al.*, 1996). The total amount of transfected DNA was fixed at 15 μg, whereas the ratio of input DNAs was varied. Excess L over H chain DNA was used in attempt to avoid the L chain from being limiting. Formation of heterodimers and homodimers was assessed by protein A purification followed by SDS-PAGE, staining with Coomassie brilliant blue R-250 and scanning laser densitometry (Ridgway *et al.*, 1996).

Chapter 4. Functional Selection for Catalysis

INTRODUCTION

Enzymes catalyze a variety of biological reactions, making them extremely attractive targets for protein engineering. There are two basic approaches to engineering enzymes. One is a knowledge-based approach where one often generates several designed mutants by modifying specific residues on the enzyme (for example Carter & Wells, 1987; Abrahmsen *et al.*, 1991; Quemeneur *et al.*, 1998). Another approach is random mutagenesis followed by screening or selection, which offers the tremendous advantage that many variants can be assayed quickly, typically 10^3 to 10^7 (for recent reviews see Kast & Hilvert, 1997; Zhao & Arnold, 1997). However, mutant proteins are tested in a biological setting, which limits the range of properties that can be engineered and complicates assays. Thus, a general and robust *in vitro* selection method is highly desirable.

Phage display is an extremely powerful *in vitro* selection technique where $>10^9$ protein or peptide variants can be subjected to selection for improved binding properties (for reviews see Clackson & Wells, 1994; Wilson & Finlay, 1998). Modest improvements in enzyme and catalytic antibody function have been obtained by selecting for binding to transition-state analogs (Hansson *et al.*, 1997; Baca *et al.*, 1997; Fujii *et al.*, 1998; Arkin & Wells, 1998). However, transition state analog binding did not correlate with improved catalysis (Baca *et al.*, 1997; Arkin & Wells, 1998). Specialized selections using reactive substrates (Janda *et al.*, 1994), inhibitors (Licht & Lerner, 1995), active site ligands (Widersten & Mannervik, 1995), and reactive products (Janda *et al.*, 1997) have also been used in attempts to select for improved catalysts. A more general functional selection has been developed by investigators interested in novel catalytic RNAs (Larsch & Szostak, 1994). Catalytic RNA molecules have been efficiently evolved for ligase-type reactions

using a product capture approach wherein the reaction modifies the RNA molecule itself and allows for affinity capture (Zhang & Cech, 1998).

We wished to develop such a product capture selection using phage display. It was desirable to improve subtiligase, a variant of the bacterial serine protease subtilisin BPN', that was engineered to carry out the ligation of peptides (Abrahmsen *et al.*, 1991; see also Holford & Muir, 1998; Dawson *et al.*, 1994; Bongers & Heimer, 1993). Subtiligase catalyzes the ligation of C-terminal activated peptides to the N-terminal α-amine of peptide and protein acceptors (Figure 4-1A) and has been used to synthesize proteins with unnatural amino acids (reviewed in Braisted *et al.*, 1997). We reasoned that its catalytic efficiency might be improved by requiring that subtiligase on phage catalyze the attachment of a biotinylated-synthetic peptide onto its own N-terminus (Figure 4-1B). The biotinylated phage could be captured efficiently by binding to immobilized neutravidin.

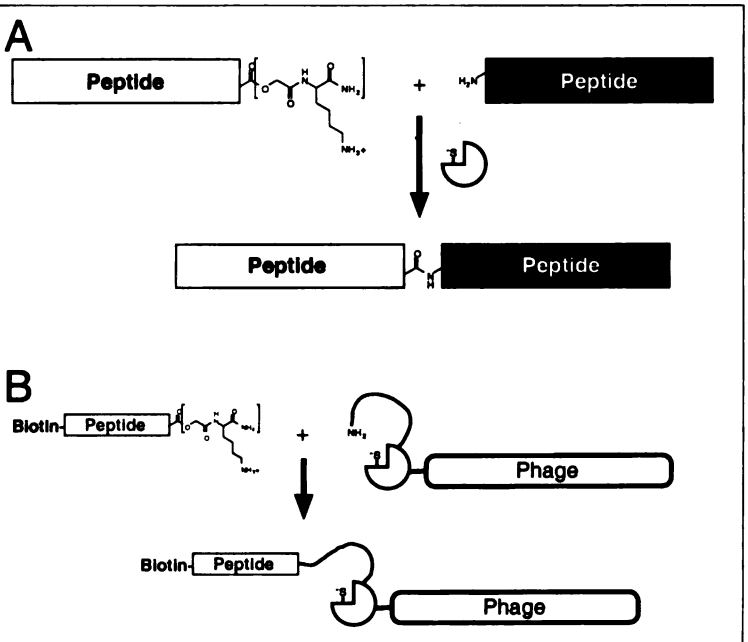

Figure 4-1. (A) Subtiligase catalyzed ligation of peptides in solution. (B) Scheme for selecting active subtiligase mutants on phage by requiring that the enzyme attach an activated peptide onto its extended N-terminus. Biotin labeled subtiligase-phage are then captured with neutravidin.

Here, we report the successful implementation of this product capture strategy. The results of the selection recapitulate and extend many years of research on subtilisin (reviewed in Wells & Estell, 1988; Perona & Craik, 1995). As expected, the active site residues, including the catalytic triad and oxyanion hole, were highly conserved in the selection. Two new mutants were identified that increased the activity of subtiligase. Surprisingly, many other mutants were selected that increased

stability, resistance to oxidation and display on phage. All of these effects increase functional display. This is a powerful strategy that directs simultaneous selection for a number of improved functional properties.

RESULTS

Phage display of functional subtiligase. The challenge in developing a product capture strategy was to demonstrate functional display of a subtiligase variant that would ligate peptides onto its own amino terminus. Subtilisin BPN' is naturally expressed as a pre-pro-protein. The pre-sequence is removed upon secretion. The pro-domain participates in folding the enzyme and is then removed in an autocatalytic process. The enzyme has a high affinity calcium binding site. Deletion of the calcium binding loop destabilizes the enzyme but allows it to fold reversibly and independently of the pro-domain (Bryan *et al.*, 1992).

A phage ELISA was used to test for display of subtiligase on phage. Phage from serial dilutions of concentrated stocks (10^{13} ml^{-1}) were captured using an immobilized anti-subtilisin antibody and detected with horseradish peroxidase conjugated to anti-phage antibodies. Using this phage ELISA we could detect display of subtiligase on phage when it was expressed under the control of a P_{tac} promoter as a pre-pro-protein fused to gIIIp. Removing the calcium loop residues 75 to 83 (ΔCa) improved subtiligase display 350-fold, possibly by allowing subtiligase to fold more reversibly (Figure 4-2). Co-expressing the pro-domain (Δpro + pro) had little effect, but removing the pro-domain altogether (Δpro) abolished display. Combining the Δpro and ΔCa mutations allowed the enzyme to fold to its mature form without the assistance of the pro-domain and without the requirement that subtiligase autoproteolyze the pro-domain. We therefore chose this construct for selection experiments.

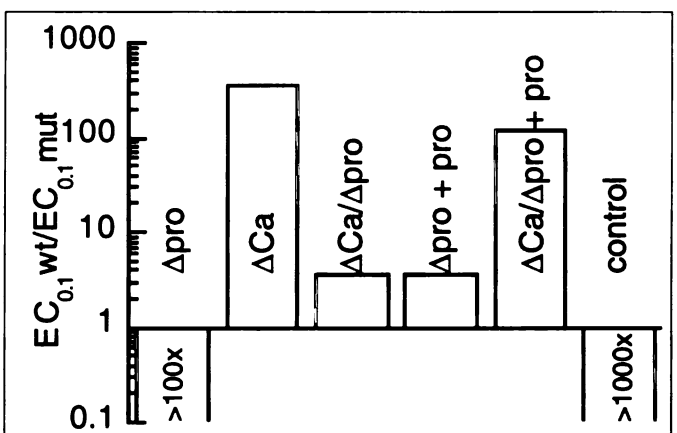

Figure 4-2. Improved display of engineered subtiligase. Serially diluted phage were captured with an anti-subtilisin antibody in a phage ELISA. Bars indicate how much less mutant phage versus wild-type was needed to generate the same ELISA signal. Phage were quantified by OD_{269}.

In order to demonstrate the catalytic functionality of the displayed subtiligase we utilized a covalent serine protease inhibitor, biotin-AAF-chloromethylketone, that reacts with the active site cysteine of thiol-subtilisin (S221C) (Phillip & Bender, 1983). Subtiligase phage that were reacted with the biotin-inhibitor were efficiently captured on neutravidin plates in proportion to their display on phage (data not shown). Incubating the phage with 1 mM DTT to reduce the active site cysteine lead to a 100-fold increase in capture by the inhibitor (data not shown). Like cysteine proteases (Lowe, 1976), subtiligase is prone to oxidative inactivation. We speculated that the reduction eliminated a mixed disulfide between the active site cysteine and a small molecule thiol. Thus, selections, ELISA's and kinetic experiments were performed with 1 mM DTT.

Subtiligase-phage catalyzed self-ligation. The strategy for selecting active enzymes requires that the N-terminus reach the active site to function as a nucleophile. The N-terminus of mature subtilisin is ~30 Å from the P1 recognition site (Bott *et al.*, 1988; Gallagher *et al.*, 1995). Structural considerations and the behavior of unstructured peptides (Creighton, 1993) suggested that the N-terminus needed to be extended by 10 to 15 residues to reach the active site. We therefore constructed several derivatives of ΔCa subtiligase with increasing N-terminal extension lengths. There was a dramatic improvement in ligation efficiency after a six residue extension (Figure 4-3A). Although the effective concentration of the N-terminus near the active site is probably in the 10 to 100 mM range (Creighton, 1993), it was important to show that ligation was occurring intramolecularly. We determined that subtiligase-phage does not ligate a peptide onto the N-terminal extension of an inactive enzyme, demonstrating that the ligation is intramolecular (Figure 4-3B).

Previous studies have shown that efficient ligation depends on the sequence of the nucleophilic peptide (Chang *et al.*, 1994). Thus it was desirable to optimize the residues of the N-terminal nucleophile on subtiligase. We constructed a subtiligase-phage library with a 15 residue extension of the form $XXX(SEGGG)_2XX$. This library had three randomized residues at the N-terminus, followed by a 10 residue flexible linker and two more randomized codons joined to the natural N-terminus of subtiligase. In order to avoid folding and autoprocessing artifacts the Δpro/ΔCa construct was used and the extension was inserted between the signal sequence and the natural N-terminus. The library was sorted by capturing the ligation product after incubating the subtiligase-phage with biotin-peptide substrate. Phage sorting was monitored using the same reaction and measuring captured phage by ELISA. After three rounds of sorting there was little further increase in captured phage and therefore functional selection was largely complete.

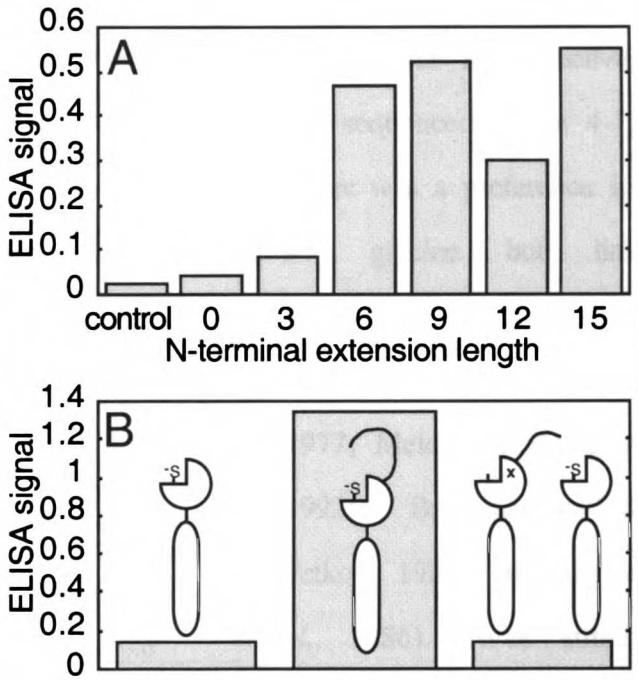

Figure 4-3. (A) Capture of subtiligase with various extension lengths. 10 µl of concentrated phage (approximately 10^{13} /ml) were incubated with 100 nM biotin-SGGKAAPFglcK-NH_2 in PBS, 10% DMSO before diluting tenfold with blocking buffer, capturing with neutravidin, and determining an ELISA signal. Because the ELISA signals are not normalized for phage concentrations or display levels, the drop in activity for the 12-mer extension might be due to factors other than substrate availability. The same reaction was used to determine (B) intra- versus intermolecular ligation, using subtiligase-phage, subtiligase-phage with an extension, and subtiligase-phage mixed with inactive subtiligase phage with an extension.

Table 4-1. N-terminal extension selectants in the Δpro/ΔCa library.

signal-XXX(SEGGG)$_2$XX-ligaseΔCa						
clone	X1[a]	X2	X3...X14		X15	ELISA[b]
45	gcg A	tag *[c]	acg T	acc T	gag E	2.0
42	agg R	aag K	gag E	agc S	ccg P	1.5
43	ggg G	cag Q	gcc A	cgg R	aac N	1.2
41 (2/1[d])	aac N	aag K	agg R	gag E	tag *	1.1±0.1[e]
46	ggc G	ctg L	agg R	gac D	agg R	1.0
37	tag *	acc T	tcg S	aac N	atg M	0.9
44	tag *	gac D	agg R	agg R	ccg P	0.6
39	tag *	ctc L	gtg V	ggc G	aac N	0.6
52	gcc A	gtg V	ggg G	aag K	cag Q	0.6
54	gcc A	gac D	tag *	ggg G	ggg G	0.5
51	ggg G	ggg G	ccg P	ctg L	tcc S	0.5
40	ggg G	agc S	cgg R	cac H	gac D	0.5
47	agg R	ggg G	gcc A	cag Q	cag Q	0.3
53	cag Q	aag K	tcc S	tac Y	cgg R	0.3
50	tcg S	gcc A	atg M	gtg V	ctc L	0.2

[a]randomized codons are indicated by X followed by the residue number. DNA sequence is listed in lowercase and amino acid sequence below in upper case by single letter code.

[b]signal of identically prepared phage in a product capture ELISA, incubating with biotin-substrate and capturing with neutravidin, uncorrected for phage yield (see Materials and Methods). Only clones with a signal higher than the signal without biotin-substrate are listed.

[c]Amber stop codons are suppressed by Gln in XL1-Blue cells.

[d]Numbers in parentheses indicate the number of clones with the same amino acid sequence followed by the number of different DNA sequences.

[e]the variation in product capture ELISA signal between independently prepared phage stocks.

Sixteen clones were assayed for ligase activity and sequenced (Table 4-1). There was a preference for P1' glycine but little preference for other subsite residues (Morihara & Oka, 1977; Meldal & Breddam, 1992; Bratovanova & Petkov, 1987; Bromme *et al.*, 1986). Three similar libraries were constructed based on the full-length ligase construct (Table 4-2), the ΔCa construct (Table 4-3), and the Δpro/ΔCa + pro construct (Table 4-4). Sorting results strongly supported the P1' preference for glycine. Among the four libraries there was a prevalence of amber stop codons, rare AGG codons (Kane, 1995), positively charged residues (Boyd & Beckwith, 1990), and whole codon and frameshift deletions that nevertheless yielded excellent display levels (Jacobsson

& Frykberg, 1995; Carcamo *et al.*, 1998). These were probably selected to reduce the host toxicity of subtiligase by reducing protein expression. *E. coli* bearing the subtiligase phagemid grow slower and are more prone to lysis (data not shown). Mutations that decrease subtiligase toxicity and therefore increase culture density and phage production while still yielding the minute amounts of protein needed for incorporation into the phage coat probably lead to an overall increase in phage that display subtiligase.

Table 4-2. N-terminal extension selectants in the full length subtiligase library.

signal-pro-XXX(SEGGG)2XX-ligase						
clone	X1	X2	X3...X14		X15	ELISA
13	ggc G	aac N	atg M	tag *	ggc G	1.1
7	ggg G	gac D	cgc R	tag *	gag E	1.1
18	ggg G	tag *	atg M	ccg P	acc T	1.0
5	ggc G	ggg G	tac Y	acc T	gac D	0.9
16	ggc G	agg R	acg T	ggg G	gac D	0.8
14	ggg G	gtc V	tgg W	gag E	gag E	0.6
8	ggc G	aag K	ggg G	ttc F	ggg G	0.5
17	ggt G	tcc S	agc S	agg R	ggg G	0.4
10	ggg G	acc T	aac N	cgg R	cac H	0.4
15	agc S	gcg A	ccg P	ttg L	gtc V	0.3
12	atg M	agc S	agc S	aggc G	gag E	0.3
11	aag K	gcg A	aac N	tcc S	tcc S	0.1

[a]the three preceeding glycines were replaced by $E^{gag}A^{gcg}$ in this clone. See table 4-1 for further explanation of symbols.

We chose the Δpro/ΔCa construct with the N-terminal GLR sequence (Table 4-1 #46) because it gave an ELISA signal twofold better than the designed 15mer extension and its P1' and P2' matched the subsite preferences of subtilisin in solution experiments (Gron *et al.*, 1992). The Δpro/ΔCa construct having an optimized N-terminal extension satisfied our criteria of a subtiligase that avoided the need for pro-domain assisted folding and autoproteolysis, and catalyzed intramolecular ligation.

Table 4-3. N-terminal extension selectants in the ΔCa library.

signal-pro-XXX(SEGGG)$_2$XX-ligaseΔCa											
clone	gca[a] A-4	cat H-3	gcg A-2	tac Y-1	nns X1	nns X2	nns X3	tct S4	nns X14	nns X15	ELISA[b]
20	.	.	gcc A	gag E	ggc G	aag K	-[c]	.	gag E	cgc R	6.2
24	.	.	gta V	cga R	agc S	tgg W	-	.	gac D	aac N	6.1
19	.	.	.	tct S	aga R	gca A	-	.	agc S	ccc P	5.7
27	ca fs[d]	.	.	.	atg M	act T	-	-	gag E	ccc P	5.5
32	.	.	gct A	aca T	ggc G	ggg G	-	.	ggg G	gcc A	5.1
29	.	.	.	aca T	gga G	gca A	ac fs[d]	.	cag Q	gcg A	4.5
31[e]	.	.	.	.	gcg A[f]	ttg L	aac N	.	acc T	aag K	3.7
26	.	.	.	tag *	ggg G	cgg R	-	.	gcc A	gag E	3.4
21	.	.	.	tgg W	aac N	gtg V	-	.	aac N	tac Y	3.3
34	ca fs[d]	.	.	.	aag K[f]	gtc V	ccg P	.	aac N	agg R	3.3
25	gta V	gat D	.	.	ggc G	aag K	gcc A	.	ggc G	gac[g] D	3.2
30	.	.	.	agc S	atg M[f]	gtg V	-	.	atg M	cgg R	2.9
28	.	ctg L	cgt R	acc T	ggc G	ggg G	gct A	-	cag Q	gcg A	2.7
23	.	.	.	aac N	gtg V	ccc P	ttg L	agg[g] R			2.3
33	.	ctg L	cgt R	aca T	ggg G	gcg A	-	.	aac N	agc S	2.3
22	.	.	.	tcc S	ggg G[f]	agc S	agg R	.	agg R	agg R	1.7
35[e,h]	.	.	.	.	ggc G	ggg G	agg R	.	agg R	acc T	0.7
36	.	ctg L	cgt R	acg T	ggg G	aca T	gg fs[d]	.	agc S	ccc P	0.3

[a]spurious mutations were observed outside the randomized residues among these selectants, so the four residues preceeding and one following the first randomized residues are listed with the DNA sequence (where n=a,g,c,t and s=g,c).

[b]the product capture ELISA signal of these clones was quenched 3-fold sooner than for the other libraries (Table 4-1, 4-2, 4-3), so the signal was multiplied by 3 for comparison.

[c]dash indicates a whole codon deletion.

[d]frame shift mutations (all single base deletions) appeared in several clones.

[e]only two clones had no spurious mutations, #31 and #35.

[f]the N-terminal sequence was verified for the mature form of these proteins expressed in 33B6 cells (induced with IPTG for 2 hours at midlog phase and whole cell samples analyzed and amino acid sequenced from SDS-PAGE). The N-terminal sequence implied processing after the -1 residue as listed in this table, despite two cases of a non-optimal serine at this position and one case of a frame shift before.

[g]the continuation of these sequences did not correspond to the subtiligase sequence, but was not sequenced far enough to determine the nature of the recombination.

[h]this mutant was subcloned into pLA1019 between the BclI and BsiWI sites (pSA0984) and expressed in *B. subtilis* (see Materials and Methods). The N-terminal extension remained intact during purification when PMSF was included to inhibit other proteases.

Table 4-4[a]. N-terminal extension selectants in the Δpro/ΔCa +pro library.

signal-XXX(SEGGG)$_2$XX-ligaseΔCa + pro						
clone	X1	X2	X3...X14		X15	ELISA
64	ggg G	ttg L	cag Q	tag *	agg R	0.8
65	ggg G	ggc G	acg T	ggc G	agc S	0.5
67	aag K	agc S	gag E	agg R	tgc C	0.5
56	ggc G	atc I	tag *	atc I	gcg A	0.4
63	ggg G	tcc S	gag E	ccg P	gcg A	0.4
69	atg M	acg T	acc T	gtc V	tag *	0.3
57	aag K	ttg L	tag *	cag Q	gtc V	0.3
58	ggc G	aac N	cag Q	ccc P	gcg A	0.2

[a]see tables 4-1 and 4-3 for explanation of symbols.

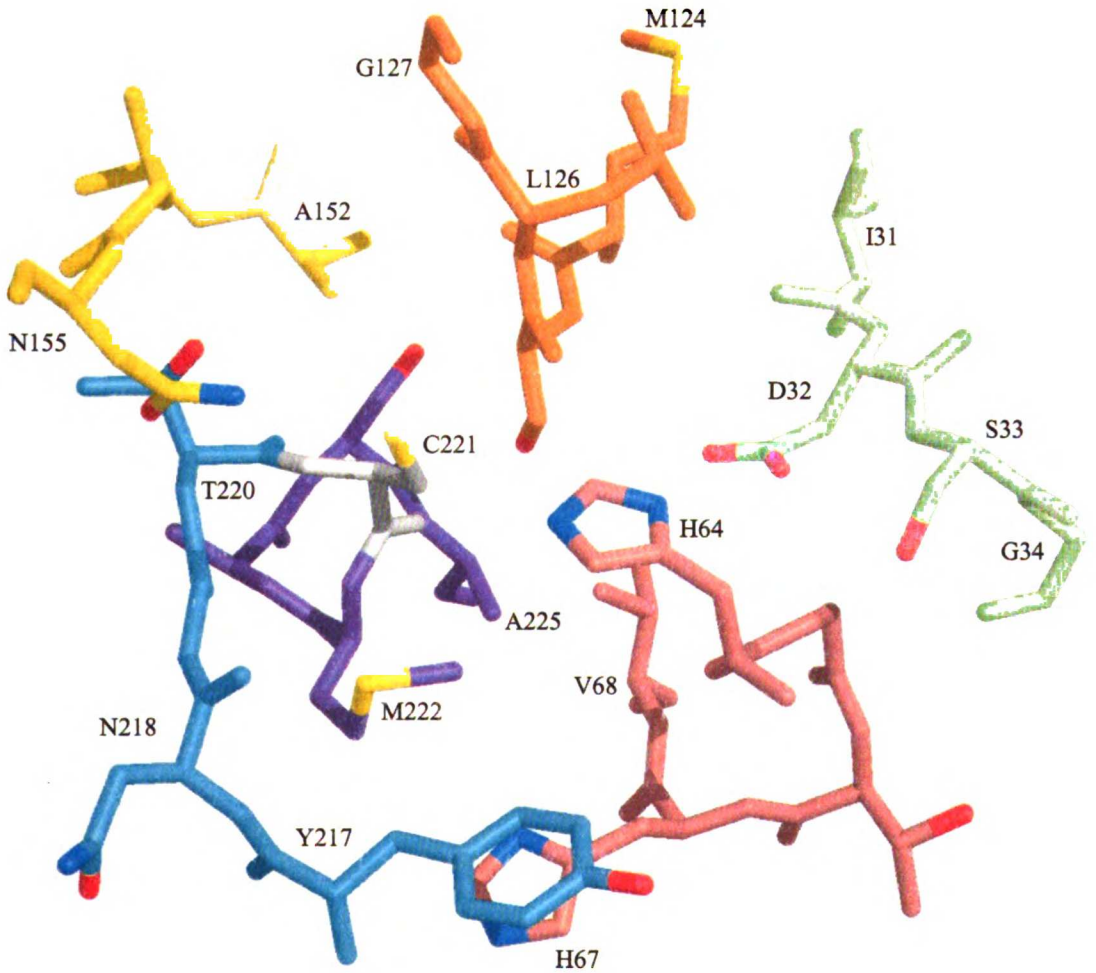

Figure 4–4. Subtiligase active site with residues randomized in the six libraries colored differently. Side–chain N, O and S are colored blue, red and yellow. Cys221 is colored white. Sulfur oxides of Cys221 are not shown. Coordinates supplied by Abraham de Vos (Abrahmsen *et al.*, 1991).

Selection for improved catalysis and characterization of mutants. Subtiligase is a designed double mutant of subtilisin BPN'. Mutating Ser221 to Cys in subtilisin leads to a 300-fold preference for aminolysis over hydrolysis (Natatsuka *et al.*, 1987); further mutating Pro225 to Ala increases turnover of ester substrates 10-fold while retaining a 25-fold preference for aminolysis (Abrahmsen *et al.*, 1991).

To optimize the subtiligase active site for ligation we targeted the 25 active site residues nearest the catalytic triad and scissile bond (Figure 4-4) including Pro225, His64 and Asp32. Four to five contiguous residues were randomized in six libraries: residues 31 to 34, 64 to 68, 124 to 127, 152 to 155, 217 to 220, and 222 to 225.

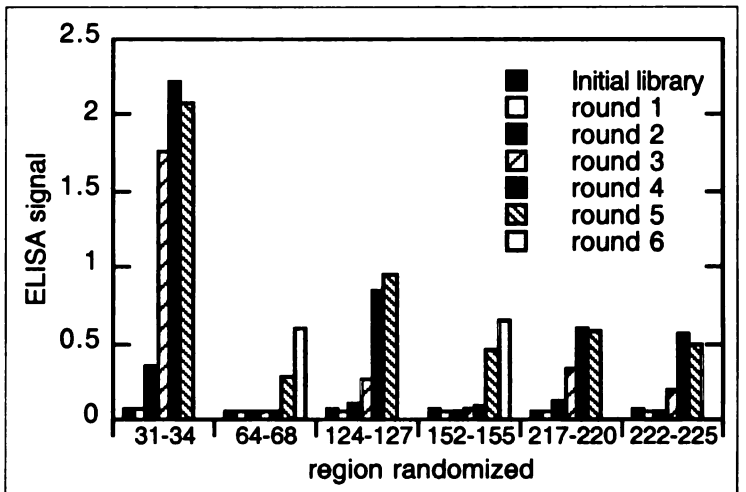

Figure 4-5. Progress of library selection monitored by product capture ELISA. Libraries were stored at 4 °C and assayed together.

The libraries were sorted separately for their ability to self-ligate a biotin-substrate until the yield of captured phage in a product capture ELISA stopped increasing (Fig 4-5). Ten to twenty clones from each library were assayed and sequenced (Table 4-5). We further analyzed the eight mutants most active on phage and I31L/S33T after expressing them as soluble proteins.

Table 4-5. Active site library selectants.

Clone	Residues randomized						increased display[a]	increased activity[b]
Library 1	I31		D32		S33	G34		
1.1 (17/4[c])	ttg ctg[d] L		gac .[e]		acg T	ggc ggg .	11	1.0
Library 2	H64	G65		T66	H67	V68		
2.1 (5/1)	cac .	ggg .		acc .	tac Y	gtg .	1.5	0.13
2.2 (4/3)	cac .	ggc ggg .		acc acg .	gcg gcc A	gtg .	3.2±0.2[f]	0.13±0.003
2.3	cac .	ggg .		acc .	tag *	gtg .	2.1	0.065
Library 3	M124	S125		L126	G127			
3.1	ctg L	tcg .		ggg G	ggg .		[g]	
3.2	ctg L	tcg .		gtc V	ggc .		1.1	2.3
3.3	ttg L	gcg A		ttg .	ggg .		0.53	2.3
3.4	ctg L	ggg G		gcg A	ggc .		0.69	0.41
3.5	ttg L	ggg G		tgc C	ggg .		2.9	1.2
3.6	atg .	ggc G		tgc C	ggg .		3.0	1.0
3.7	atc I	gcg A		gtg V	ggc .		2.7	0.48
3.8	atc I	tcg .		ttg .	ggg .			
Library 4	A152	A153		G154	N155			
4.1 (14/1)	tgc C	gcc .		ggg .	aac .		3.1	0.068
4.2 (2/1)	tcc S	gcg .		ggg .	aac .		0.40	0.079

Library 5	Y217	N218	G219	T220	increased display[a]	increased activity[b]
5.1	tac .	tcc S	ggg .	acg .	2.6	0.63
5.2	gcg A	agc S	ggg .	acg .	3.4	0.72
5.3	tgc C	acg T	ggg .	acc .	3.4	1.9
5.4	gtg V	acg T	ggg .	acg .	1.7	0.58
5.5	atg M	gcg A	ggc .	acc .	2.7	0.19
5.6	ggg G	cac H	ggg .	acc .	2.2	0.61
Library 6	M222	A223	S224	A225		
6.1 (2/2)	ggg G	gcc .	tcc tcg .	gcc .	0.49	1.9
6.2	agc S	gcc .	tcg .	gcg .	0.74	1.9
6.3	agc S	gcg .	gcc A	gcg .	0.49	1.7
6.4	gcc A	gcg .	agc .	ggc G	0.46	1.4
6.5	gtc V	gcg .	tcc .	gcc .	0.29	0.19
6.6	tag *	gcc .	tcc .	gcg .	1.5	0.14
6.7	ttc F	gcg .	agc .	gcg .	0.17	2.7
6.8	atg .	agc S	tgc C	gcg .	0.31	0.79

[a]Increased display is measured as the $EC_{0.2}$wt/$EC_{0.2}$mut in the antibody capture ELISA (see Materials and Methods). Only clones that showed activity in the product capture ELISA were sequenced and further characterized, corresponding to 17 of 31 clones for the 31-34 library, 13 of 18 for the 64-68 library, 8 of 12 for the 124-127 library, 16 of 18 for the 152-155 library, 8 of 12 for the 217-220 library and 10 of 12 for the 222-225 library. Several clones gave no or unreadable sequence.

[b]The improved activity on phage ($EC_{0.2}$wt/$EC_{0.2}$mut in the product capture ELISA) is normalized for improved display ($EC_{0.2}$wt/$EC_{0.2}$mut in antibody capture) (see Materials and Methods).

[c]numbers in parentheses indicate the number of instances of a particular amino acid sequence followed by the number of different DNA sequences for each.

[d]all the codons observed for a particular amino acid are listed.

*amber stops are suppressed with Gln in XL1-Blue cells.

[e]period indicates conservation of the parent subtiligase residue.

[f] Two clones were determined to be identical in amino acid sequence, though not in DNA sequence, after sequencing. The variability in these clones and variability in other experiments indicates that the activities on phage are reproducible to ±10%.

[g]not measured

Two methods were used to measure the ligation reaction in solution with purified protein. First, the appearance of product was monitored with HPLC using a the same substrate used for the selection (300 μM biotin-SGGKAAPFglcK-NH_2) and a tetrapeptide with the same first three residues as the N-terminal extension of the subtiligase (3 mM H-GLRY-NH_2, 1 μM enzyme) (Figure 4-6). The acceleration of the reaction with each mutant compared to the reaction with the wild-type enzyme is shown in Figure 4-7.

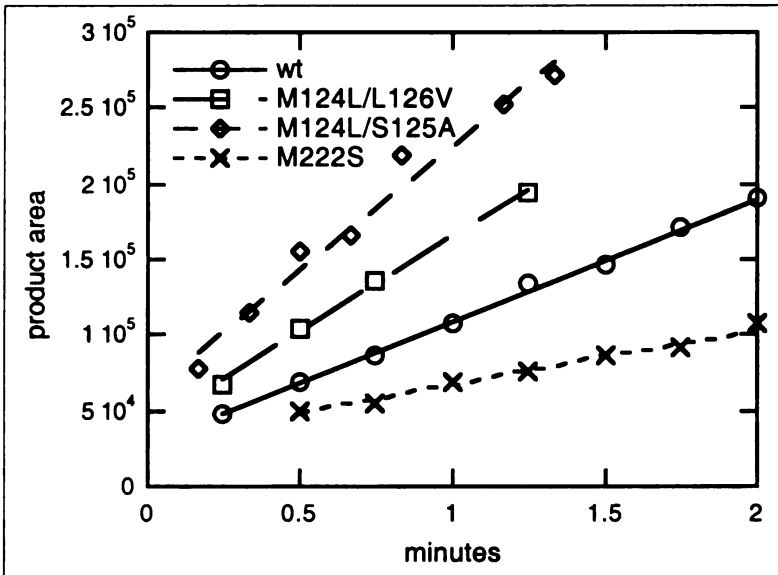

Figure 4-6. Example HPLC monitored ligation reactions.

It was difficult to detect the ligation product at lower concentrations with HPLC. So a fluorescence donor-quencher pair of substrates was synthesized in order to monitor ligation in a fluorescence quenching assay (Meldal & Breddam, 1991). *O*-Aminobenzoyl-AAPFglcK-NH_2 is excited at 320 nm and fluoresces at 407 nm, but the ligation product, aminobenzoyl-AAPFGY(NO_2)-NH_2, is 95% less fluorescent (data not shown). The initial velocity of quenching was measured for the wild-type enzyme and various mutants (40 μM ABz-AAPFglcK-NH_2/ 1 mM H-GY(NO_2)-NH_2/ 5 μM enzyme). There is good agreement between the HPLC and fluorometric assays (Figure 4-7).

There was only a 60 to 70% correlation between the increased activity measured on phage and the increased activity of the purified proteins measured in the fluorometric and HPLC assays. This could be due to the different concentrations of substrates used, the deletion of the calcium loop, or other differences in assay conditions. Since the fluorescence assay used a substrate concentration (40 μM) far below the K_M of the parent subtiligase for the activated peptide (about 1 mM (Abrahmsen *et al.*, 1991)), it is a better measure of the change in k_{cat}/K_M (Fresht, 1985; Bender *et al.*, 1964). From these data it was clear that M124L/S125A and M124L/L126V are more than twofold improved over the parent subtiligase.

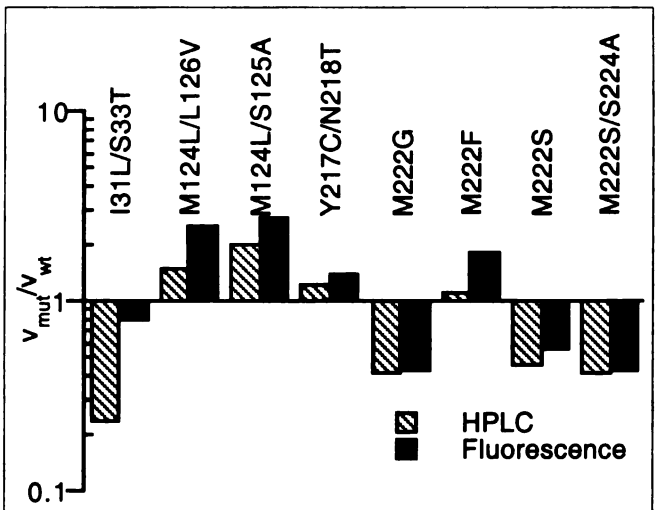

Figure 4-7. Kinetics of mutants measured by HPLC and fluorescence quenching. The improvement versus wild-type is plotted on a log scale. In the HPLC determination (300 μM biotin-SGGKAAPFglcK-NH_2 / 3 mM H-GLRY-NH_2 / 1 μM enzyme) the initial velocity of product formation catalyzed by the parent subtiligase, $v_{w.t.}$, was 58 μM s^{-1}. In the fluorescence assay (40 μM Abz-AAPFglcK-NH_2 / 1 mM H-GY(NO_2)-NH_2 / 5 μM enzyme), $v_{w.t.}$ = 0.14 μM s^{-1}. Rates were reproducible to ±10%.

DISCUSSION

Here we developed a product capture strategy for selecting subtiligase variants displayed on phage. The selection results recapitulate and expand many mutagenic studies of subtilisin BPN' (Wells & Estell, 1988). Fifteen of the twenty-five residues randomized were strongly or completely conserved as wild-type residues (Table 4-5) (Siezen & Leunissen, 1997). Asp32 and His64 of the catalytic triad were completely conserved, as was Asn155 and Thr220 of the oxyanion hole (Wells & Estell, 1988; Braxton & Wells, 1991). P225A was largely conserved with P225G the only exception, validating the original design of subtiligase.

Two mutants in the 124-127 library were significantly more active than the parent subtiligase, M124L/L126V (2.5-fold improved) and M124L/S125A (2.7-fold). M124L was possibly selected for oxidation resistance (see below); however oxidation of Met124 in subtilisin has little effect on catalysis (Bott *et al.*, 1988; Lilova *et al.*, 1987), and it is more likely that M124L, common to both catalytically improved variants, was selected for catalysis.

Some of the selected residues are specific to subtiligase as compared to subtilisin BPN'. Ser125 forms a hydrogen bond with residue 221 and is conserved throughout the subtilase superfamily (Siezen & Leunissen, 1997), but is not required in the context of subtiligase, perhaps due to less stringent nucleophile positioning requirements for esterase activity. Leu126 is conserved among subtilisins and is important for subtilisin activity (Rheinnecker *et al.*, 1994; Sorensen *et al.*, 1993), but was not conserved in the subtiligase selection. It is interesting that two mutations that are beneficial to subtiligase but harmful to subtilisin, L126V and M222F (Estell *et al.*, 1985), are on opposite sides of Cys221 (Figure 4-4). It is possible that these mutations reposition the substrate relative to the active site Cys221. These subtiligase specific mutations underscore the mechanistic differences between this engineered enzyme and wild-type subtilisin.

Some of the mutants appear to have been selected on the basis of improved thermodynamic or oxidative stability. N218S stabilizes subtilisin to heat denaturation (Bryan *et al.*, 1992). In our selection on subtiligase Asn218 was mutated largely to Thr or Ser. Met222 is a known site of oxidative inactivation. Though conserved in subtilisins, in the subtiligase library it was mostly mutated to small amino acids, known to be the least deleterious to subtilisin activity (Estell *et al.*, 1985) (Table 4.5). Nearby His67 was completely mutated to Tyr or Ala, possibly due to some role in Met222 oxidation. It is interesting to note that in model peptides a proximal histidine can accelerate the oxidation of methionine (Schoneich *et al.*, 1993).

In some cases disulfides were apparently selected. Ala152 was mutated almost completely to cysteine. This residue is close to the active site cysteine, 5.0 Å Cα to Cα using the subtiligase structure. It is possible that A152C was selected to form a stabilizing disulfide bond with the active site sulfhydril, which upon reduction generates a free Cys221 and therefore an active enzyme (Lowe, 1976; Korodi *et al.*, 1986). There are several other cysteines that appeared among the selectants: L126C (6.3 Å), S224C (5.2 Å) and Y217C (9.3 Å). Except for Y217C, the distances of these residues are within the 3.8 to 6.8 Å range of observed disulfide distances (Srinivasan *et al.*, 1990). These disulfides probably contribute to folding or stability, since most of the mutants display better on phage (Table 4-5). Besides being used to prevent irreversible inactivation and as a reductive switch, these cysteines might be utilized in the redox cycle of selenosubtilisin to assist in the reduction of peroxides (Peterson & Hilvert, 1995).

Some variants were selected to increase display on phage. I31L/S33T was completely conserved in the 31-34 library. This double mutant displays 10-fold better on phage than wild-type, and is approximately equal to the wild-type catalytically (Table 4-5, Figure 4-7). The selected double mutation might be beneficial to subtiligase display by facilitating folding, by analogy to the beneficial effect of removing the Ca loop, since I31L is known to decrease the stability of subtilisin E (Takagi *et al.*, 1988).

Improving enzyme function using phage display has been a long sought goal. Previous selections (Hansson *et al.*, 1997; Baca *et al.*, 1997; Fujii *et al.*, 1998; Arkin & Wells, 1998; Baca *et al.*, 1997; Arkin & Wells, 1998; Janda *et al.*, 1994; Licht & Lerner, 1995; Widersten & Mannervik, 1995; Janda *et al.*, 1997) and attempts to develop selections (Corey *et al.*, 1993; Maenaka *et al.*, 1996; Soumillion *et al.*, 1994a; Gao *et al.*, 1997; Soumillion *et al.*, 1994b; Pederson *et al.*, 1998) for catalysts on phage have largely met with difficulty. Here we have developed a product capture strategy to select improved subtiligase mutants on phage. Several mutants were selected with improved catalytic activity. Surprisingly, many others seem to be improved in other characteristics, including stability, folding, and oxidative resistance. These other mutations were likely selected because they contribute to functional display on phage. If it was desirable, selection for these other characteristics could be avoided by altering library construction and selection conditions. Alternately, one could specifically focus selection pressure on a particular characteristic, for example, by sorting in the presence of denaturants or oxidants.

Due to the wide range of mutations identified, the product capture selection has confirmed and expanded long-studied subtilisin features in to the context of subtiligase: conservation of active site residues, stability mutants, oxidation resistant mutants, new disulfide bonds, and folding mutants. We believe that the product capture strategy can be successfully applied to other enzymes with peptide substrates and to enzymes with non-peptide substrates by the use of appropriate substrate and product capturing reagents (Tawfik *et at.*, 1993; Pederson *et al.*, 1998; Demartis *et al.*, 1999; Jestin *et al.*, 1999).

MATERIALS AND METHODS

Synthesis of Peptides.

The subtiligase inhibitor biotin-AAF-chloromethylketone (Philipp & Bender, 1983) was synthesized in 360 µl of 18 mg/ml H-AAFcmk (Bachem), 40 mg/ml sulfo-NHS-LC-Biotin (Pierce), 50% DMSO, 50 mM Tricine (pH 8.0) for 1 hr. The product was purified on a preparative C18 HPLC column. The substrate biotin-SGGKAAPFglcK-NH_2 was supplied by Andrew Braisted. Abz-AAPFglcK-NH_2 was synthesized according to (Braisted & Wells, 1997) with Fmoc chemistry and N-Boc-o-aminobenzoic acid (Bachem). H-GY(NO_2)-NH_2 was synthesized from Boc-Gly-OH (1 mmol) and 3-nitro-L-Tyrosine ethyl ester (1.1 mmol) (Sigma) in DMA (3 ml) with 2.0 mmol triethylamine using HBTU (1.1 mmol). After 30 minutes the reaction was filtered, diluted 20-fold with benzene, extracted with brine and 5% HCl and concentrated *in vacuo*. The oil was dissolved in methanol (100 mls) and ammonolysed overnight with ammonia to Boc-GY(NO_2)-NH_2 according to Bodanszky & Bodanszky (1994). The product was concentrated *in vacuo*, purified on a silica column with ether as a solvent, concentrated *in vacuo*, and the BOC group cleaved with trifluoroacetic acid (5 mls) for 1 hr. The final product was concentrated *in vacuo* and purified on a silica column with methanol as solvent.

The remaining peptides were synthesized using standard methods (Barany & Merrifield, 1980). All peptides were purified on a preparative C18 HPLC column and lyophilized. Concentrations were calculated from dry weight. Purity was monitored by HPLC and masses verified by mass spectrometry.

Construction of subtiligase display vectors. pSA08541 was constructed by subcloning a fragment of pMal-p2 (New England Biolabs) encoding the *lacI*^{q} gene, the P_{tac} promoter, and the signal peptide from maltose binding protein ("malsignal") between the EcoRI and NsiI sites of pH0753 (Atwell *et al.*, 1997) (using 5'-AAAAGAATTCCCGAC-ACCATCGAATGGTGC-3' and 5'-ACCGCCACTCTAGACCTGAGCTGCCGCCTG-3'). The pro-subtiligase gene was subcloned from pLA1019 (see below) into this vector between

the NsiI and XbaI sites (using LP#1 5'-TCCTCTGATGCATCTGCAGGGAAA-TCAAAC-3' and LP#4 5'-ACCGCCACTCTAGACCTGAGCTGCCGCCTG-3') to generate pSA0906, a subtiligase phage display vector expressing malsignal$^{1\text{-}26}$-S-pro$^{1\text{-}77}$-subtiligase$^{1\text{-}275}$-V-*-gIIIp$^{249\text{-}406}$ (where * indicates an amber stop suppressed in XL1-Blue cells, and gIIIp is the gene III protein of bacteriophage M13). Derivatives of pSA0906 lacking the pro domain (Δpro) (pSA0928) (using 5'-GTTTTCCGCCTCGGCATATGCA-GCGCAGTCCGTGCC-3'), the calcium binding loop residues 75 to 83 (ΔCa) (pSA0933) (using 5'-GCCGGCACAGTTGCCGCGGTTGCGCCAAGCGCATCAC'-3'), or both (Δpro/ΔCa) (pSA0940) were constructed by site-directed mutagenesis (Kunkel *et al.*, 1991). Further derivatives with the pro region co-expressed were constructed by subcloning the signal peptide and pro domain from pSA0906 into the NheI site of pSA0928 and pSA0940 (using 5'CTTTTGGCTAGCTTCACCAACAAGGAC-3' and 5'-GGGCGGCGCTAGCTTAGTACGCATGTGCTAC-3') to generate pSA0953 (Δpro + pro) and pSA0947 (Δpro/ΔCa + pro). Mutagenesis was verified using dideoxy sequencing with dITP (Amersham) to avoid compression artifacts.

In the course of developing the subtiligase display system several other constructs were tested by capturing with polyclonal anti-subtilisin antibody in a phage ELISA. It was found that 1) the P_{tac} promoter with malsignal produced 200-fold higher display levels of sutbilisin than the P_{phoA} promoter with the stII signal. 2) Subtilisin displayed more than subtiligase. 3) The gIIIp anchor produced much higher display levels than the gVIIIp anchor (for which display was undetectable). 4) The stII secretion sequence was equivalent to the malsignal (tested in the context of the P_{tac} promoter and subtiligase). 5) Using the P_{tac} promoter led to a decrease in phage yields and culture growth, and induction with IPTG led to further reduced phage yields and loss of the ELISA signal (unpublished results).

Subtiligase display vectors with various N-terminal extension lengths were constructed by site-directed mutagenesis of pSA0933, with the following inserts between the pro domain and the second residue of ΔCa subtiligase: GLGY, GLSEGGY, GLSEGGGSGY,

GLSEGGGSEGGGY, GLSEGGGSEGGGSEGY (using oligos 5'-GTAGCACATGCG-TACGGCTTGGGGTACCAGTCCGTGCCTTAC-3', 5'-GTAGCACATGCGTACGGC-TTGTCTGAGGGAGGGTACCAGTCCGTGCCTTAC-3', 5'-GTAGCACATGCGTAC-GGCTTGTCTGAGGGAGGCGGTTCAGGGTACCAGTCCGTGCCTTAC-3', 5'-GTA-GCACATGCGTACGGCTAGTCTGAGGGAGGCGGTTCAGAAGGAGGTGGGTAC-CAGTCCGTGCCTTAC-3' and 5'-GTAGCACATGCGTACGGCTTGTCTGAGGGAG-GCGGTTCAGAAGGAGGTGGCTCTGAAGGGTACCAGTCCGTGCCTTAC-3').

For the experiments determining whether ligation was intra- or intermolecular we used pSA0940 and clone 46 of the pSA0940 extension library (pSA0940.46) (see below) to avoid autoprocessing artifacts. A pSA0940.46 derivative with the catalytic Cys221 mutated to Ala (pSA0988) was generated by site-directed mutagenesis (using 5'-GCGTACAACG-GTACTGCAATGGCATCTGCGCACGTTGCCG-3').

Phage manipulations. Phage were quantified using a molar extinction coefficient of 10^4 per nucleotide at 269 nm (Crissman, J. W. & Smith, G. P., 1984), which corresponds to 9×10^{12} phage per ml per OD_{269} for a 7000 nt phagemid. Phage manipulations were as described (Lowman, 1998). Phage were captured for ELISAs using a polyclonal anti-subtilisin antibody or neutravidin (Pierce). To capture biotinylated subtiligase phage with neutravidin, phage in phosphate buffered saline (PBS) containing 1 mM 1,4-dithio-D,L-threitol (DTT) were incubated with biotin-inhibitor or biotin-substrate. Non-specific capture of the phage became evident above 1 μM inhibitor, so we used a low concentration of the inhibitor. Selection of N-terminal extension variants and product capture ELISAs utilized 100 nM biotin-inhibitor or 100 nM biotin-substrate (in PBS, 10% dimethylsulfoxide (DMSO)) for 30 minutes, followed by tenfold dilution with casein blocking buffer (Pierce) and capturing with neutravidin in one well. Sorting of active site libraries utilized 1 μM substrate (1% DMSO) for 30 minutes followed by precipitation of phage with 1/4 volume PEG/NaCl, resuspension in blocking buffer and capture with neutravidin in 8 wells. Bound

phage were propagated by addition of midlog phase *E. coli* XL1-Blue (Stratagene) directly to wells after washing.

Increases in display levels were determined by fitting antibody capture ELISA curves to a+(b-a)/(1+(([phage]/c)^d)*((b-0.5)/(0.5-a))), in which a, b, c and d are fitting paramaters. C is the $EC_{0.5}$, the concentration of phage needed to give an OD_{492} of 0.5. This equation can be modified to determine $EC_{0.1}$ or $EC_{0.2}$. The $EC_{0.5}$ of the wild-type was divided by the $EC_{0.5}$ of the mutant ($EC_{0.5}$wt/$EC_{0.5}$mut). Ligation activities were determined using the $EC_{0.5}$wt/$EC_{0.5}$mut determined from a product capture ELISA, and normalizing for display levels by dividing by the $EC_{0.5}$wt/$EC_{0.5}$mut from an antibody capture ELISA.

Construction of phage libraries. The N-terminal extension libraries were made by site-directed mutagenesis of pSA0906, pSA0933, pSA0940 and pSA0947, inserting XXX$(SEGGG)_2$XX between the signal peptide or pro domain and the first residue of ΔCa subtiligase (using 5'-GTAGCACATGCGTACNNSNNSNNSTCTGAGGGAGGCGGTT-CAGAAGGAGGTGGCNNSNNSGCGCAGTCCGTGCCTTAC-3' or 5'-GCCTCGGC-TTATGCANNSNNSNNSTCTGAGGGAGGCGGTTCAGAAGGAGGTGGCNNSNN-SGCGCAGTCCGTGCCTTAC-3'). The procedure used was that of Lowman (1998) as modified by Sidhu and Wells (in preparation). The total number of transformants was $4.0x10^9$.

Active site libraries were designed with the aid of a graphics workstation (Silicon Graphics), using the coordinates for subtilisin BPN' with inhibitor (2st1.pdb (Bott *et al.*, 1988)), and with the pro domain bound (1spb.pdb (Gallagher *et al.*, 1995)). Libraries were constructed with stop templates (Lowman, 1998) based on pSA0940.46 (using

5'-GTAGCGGTTTAATAATAATAATCGATTCTTC-3' and

5'-GTAGCGGTTNNSNNSNNSNNSATCGATTCTTC-3',

5'-CAACAACTCTTAATAATAATAATAGCCGGCACAG-3' and

5'-CAACAACTCTNNSNNSNNSNNSNNSGCCGGCACAG-3',

5'-CGTTATTAACTAATAATAATAGGACCTTCTG-3' and

5'-CGTTATTAACNNSNNSNNSNNSGGACCTTCTG-3',

5'-GTCGTTGCGTAATAATAATAGAAGGCACTTC-3' and

5'-GTCGTTGCGNNSNNSNNSNNSGAAGGCACTTC-3',

5'-CAAATACGGGGCGTAATAATAATATGCATGGCATC-3' and

5'-CAAATACGGGGCGNNSNNSNNSNNSTGCATGGCATC-3', and

5'-CGGTACTTGCTAATAATAATACACGTTGCCG-3' and

5'-CGGTACTTGCNNSNNSNNSNNSCACGTTGCCG-3'). The total number of transformants for the six libraries ranged from $2.3x10^9$ to $4.0x10^9$.

Construction of *Bacillus* expression vectors. The plasmid pLA1019 is a variant of the *B. subtilis* expression vector pSS5 that encodes subtiligase (subtilisin BPN' S221C/P225A) (Abrahmsen *et al.*, 1991). pSA1054 was constructed by silently removing the PvuI site in the ampicilin resistance gene using site-directed mutagenesis (using 5'-GC-TCCTTCGGTCCTCCGATGGTTGTCAGAAGTAAGTTGG-3'). pSA1052, a vector with a unique PvuI site in the region of the calcium binding loop was created by removing the PvuI site in the ampicilin resistance gene, an additional PvuI site in subtiligase (using 5'-CGGAATCGAGTGGGCCATCGCAAACAATATGGACG-3') and introducing a unique PvuI site in the calcium binding loop (using oSA10451 5'-GCCGGCACAGTTGCGGCT-CTTAATAACTCGATCGGTGTATTAGGCGTTGCGCCAAGCGCATCAC-3'), all with silent mutations by site-directed mutagenesis.

Mutations in the residue 31-34 region of the subtiligase selectants were transferred into the expression vector by subcloning the BsiWI/PvuI fragment of the mutant into pSA1052 (using LP#1 and 5'-GTGATGCGCTTGGCGCAACGCCTAATACACCGATCGAGTTA-TTAAGAGCCGCAACTGTGCCGGC-3'). Mutations in the remaining selectants were subcloned into the PvuI/BclI sites of pSA1054 (using oSA10451 and LP#4). Subcloning was verified by sequencing.

Expression and purification of subtiligase mutants. DNA (1 μg) prepared from JM101, MM294 or GM48 *E. coli* was used to transform *B. subtilis* strain BG2036 as

described (Anagnostopoulos & Spizizen, 1961). Subtiligase mutants expressed poorly in *E. coli* and proteolytic degradation was evident when expressing ΔCa mutants in *B. subtilis*. Therefore, mutations were subcloned into pLA1019. Expression and purification was as described (Abrahmsen *et al.*, 1991) except that PD10 columns (Pharmacia) were used to exchange the final buffer to 100 mM Tricine (pH 8.0), 5 mM $CaCl_2$. Subtiligase concentrations were determined by densitometric scanning of Coomassie-stained SDS gels using several concentrations of subtiligase (Cunningham *et al.*, 1989). A152C subtiligase did not express. The yield of M124L/L126V was 3-fold greater than the parent subtiligase, Y217C/N218T 8-fold less, M222F 11-fold less, and M222S/S224A 2-fold less. The other characterized mutants yielded approximately as much protein as the parent enzyme.

HPLC and fluorescence determination of kinetics. 22.5 μl aliquots of the ligation reaction: 300 μM biotin-SGGKAAPFglcK-NH_2 / 3 mM H-GLRY-NH_2 / 100 mM Tricine (pH 8.0) / 0.005% tween-20 / 1 mM DTT and 1 μM enzyme, were quenched at 15 second intervals with 2.5 μl acetic acid and stored on ice. Samples were analyzed on a C18 HPLC column using a 0-90% acetonitrile / 0.1% TFA gradient, monitoring product peak areas at 214 nm. The inclusion of DTT up to 10 mM did not catalyze product formation in the absence of enzyme, did not change the initial velocities measured for freshly thawed enzyme and led to more linear product appearance plots.

Fluorescence quenching of ABz-AAPFglcK-NH_2 by ligation to H-GY(NO_2)-NH_2 (100 μM ABz-AAPFglcK-NH_2 / 1 mM H-GY(NO_2)-NH_2 / 0.005% tween / 100 mM Tricine (pH 8.0) / 1 mM DTT / 5 mM $CaCl_2$ and 5 μM enzyme) was monitored on an SLM 8000 fluorimeter (SLM Aminco) by excitation at 333 nm and monitoring at 400 nm after normalizing the signal and measuring a baseline. Due to low yields of expressed protein, M222F and Y217C/N218T fluorescence kinetics were compared to wild-type at 1 μM enzyme. Due to intramolecular quenching at such high concentrations of quencher (Meldal & Breddam, 1991; Wu & Brand, 1994) the fluorescence signal was low. There was also a 50% residual fluorescence signal after the reaction had reached completeness (as judged by

HPLC), perhaps for the same reason. Since the reactants and the product are identical among the mutants tested, the ratio of mutant activity versus wild-type was the measurement of interest, and the kinetics agreed with the HPLC measurements, we concluded that the fluorescence assay gave an accurate measurement.

The best nucleophile for quenching was H-GY(NO_2)-NH_2, but several other peptides were tested and found to be no more than two-fold worse (as measured by the rate of quenching for the wild-type enzyme; data not shown): H-AY(NO_2)-OEt, H-GLY(NO_2)-NH_2, H-GLNY(NO_2)-NH_2, H-GLPY(NO_2)-NH_2, H-GF(NO_2)-NH_2 (Yaron & Katchalski-Katzir, 1979), and H-GLF(NO_2)-NH_2. Quenching was less than 0.01%/sec in the absence of nucleophile, in the absence of enzyme, with H-GLRY-NH_2, or with H-GLD-NH_2. The inclusion of 1 mM DTT did not significantly effect the initial velocities measured.

Chapter 5. Improving display of a cytokine receptor using DNA shuffling

INTRODUCTION

The most widespread use of phage display has been to increase the binding affinity of a given protein (reviewed in Katz, 1997) or to identify novel binders from naïve libraries (reviewed in Lowman, 1997; Griffiths & Duncan, 1998). The selection process involves capturing active protein displaying phage with an immobilized target molecule, followed by infection of bacteria with these captured phage, culturing the infected bacteria and preparing a new phage stock. Some proteins are difficult to display on phage because of problems with protein expression, secretion, folding, packaging, or stability. For these proteins it is valuable to identify the problem and particular mutations that may alleviate the problem.

The interaction between human growth hormone and its receptor mediates long-bone growth and development. Efficient display of hGH on phage has greatly facilitated the analysis and engineering of hGH, including affinity maturation of the primary binding site (Lowman and Wells, 1993), the secondary binding site (Pearce *et. al.*, 1999), and compensatory mutagenesis to recover binding to a receptor mutant (Chapter 2). However, the extracellular portion of the receptor, hGH binding protein (hGHbp), displays poorly on phage. To improve the display of hGHbp and therefore enable more extensive functional studies and engineering on the receptor side of the hGH:hGHbp interface, we decided to utilize phage display in combination with DNA shuffling (Stemmer, 1989). DNA shuffling is a powerful technique for generating and recombining random mutations in an artificial evolution system the works by fragmenting mutagenized or homologous genes and recombining them in a PCR like reaction (reviewed in Patten *et al.*, 1997).

Here we report the identification of mutants that increase functional hGHbp display on phage by greater than 60-fold. The sequences for 33 clones showed that selected mutations were concentrated in the unstructured regions of hGHbp. From the sequencing information and the behavior of individual clones in several assays it was possible to hypothesize a role for various mutations in host toxicity, folding, binding, and secondary RNA structure. These studies illustrate a straight forward approach to improve protein display and make this powerful selection tool more accessible for engineering protein properties.

RESULTS AND DISCUSSION

Initial attempts to perform DNA shuffling were hampered by low recovery of gel purified digestion fragments. Therefore we attempted shuffling without purifying the digestion fragments. We performed the DNAse fragmentation directly in PCR buffer and determined the DNAse concentration necessary to cut all of the full length material. This modification avoids the losses associated with purifying the fragmented DNA products, but does not guarantee that the shuffled gene is assembled from several different fragments.

The primerless PCR reaction itself has been found to generate a diverse pool of mutations. We found that the initial library had approximately 1 mutation per 1000 bases. Since the initial library size was 2.5 x 10^7, every point mutant and more than half of the double mutants of the 700 base gene should have been represented. One hundred point mutations were identified after sequencing clones from the initial, round 5 and round 14 libraries (Table 5-1). The observed distribution of mutations shows a 3.5-fold preference for transitions versus transversions. This is due to the additional unfavorability of mis-pairing two pyrimidines or two purines over mis-pairing a pyrimindine with a purine (Cadwell and Joyce, 1992).

Table 5-1. Distribution of mutations[a].

From\to	T:A	G:C	C:G
A:T	9	38	6
C:G	26	17	
G:C	4		

[a]mutations were sequenced from the initial library (7 mutations), round 5 (21), and round 14 (72). Mutations that appeared in multiple clones were only counted once, since they were likely derived from a unique mutagenic event. Since it is impossible to determine which strand was mutagenized during the PCR reaction, mutations are classified by basepair.

The library was sorted for binding to immobilized hGH. The enrichment, the ratio of phage bound to an hGH coated well versus an uncoated well, increased from the first to third round. On round 4 the enrichment was not improved, so DNA from the previous round was used as a template for DNA shuffling (Figure 5-1). Overall, DNA shuffling was performed three times, interspersed with three to four rounds of binding selection until a 100-fold enrichment was reached. (The "sort" or "round" is used to refer to the binding selections, with shuffling occurring before the first, fifth and tenth round of

binding selection.) Greater enrichments could probably be attained, but the 100-fold enrichment was sufficient for improving display of hGHbp to a practical level.

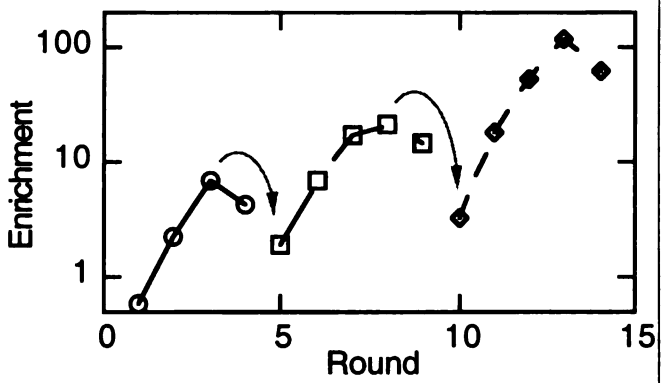

Figure 5-1. The progress of the sorting was monitored by enrichment, the ratio of phage eluted from an hGH coated well versus an uncoated well. Grey arrows indicate which pools were used for subsequent rounds of DNA shuffling and where the sorting resumed.

Table 5-2.

A. Round 5 clones

	1 - 6	7*	8 - 55	56	57	58	59	60 - 69	70	71 - 104	105	106 - 234
1						a174g						
3			g47c W16S			a174g						t486c
4					g169t G57C							
5			g94a c152t g154a E32K T51I D52N					c193g a205t Q65E T69S		t238c W80R		
6	t2c F1S					a174g						
7			c108t									g376a g430a D126N V144I
8			g22a a151g A8T T51A		g169a G57S					t298c F100L		
9	g8c G3A		c52t				a175g K59E	c206g T69S				g482c R161P
10			c144t					a186g		t226c W76R		

B. Round 14 clones

1							a175g K59E		g209c R70T				c488t A163V
2		a19c T7P		t168c				g195a					g373c V125L a554c K185T
3		c20t T7I				a172t T58S			g209c R70T	g212a a221g t231c Q74R			
4			t27c		t171c					c309t			a390g
5				t168c									t372c c402g c429g N143K
6			a106g T36A			a172t T58S							t348c g426a c669a
7			g116a R39H		g169a G57S						t314c I105T	a315c	
8							a175g K59E						
9			g126c E42D		g169a G57S								
10			g94a E32K				a175g K59E			c230g a234t T77S Q78H			a489c
11			t29c I10T					c196t c207g	G209c R70T				a574g I192V
12	g8t G3V	c20t T7I	t103c t123a F35L	t168c								t375c	
13					t171c							g387a	
14							a175g K59E				t314c I105T		
15			g71a G24D			a172t T58S		a178g N60D					t590c V197A
16		c20t T7I	a76g c107t K26E T36I			a173g				t295c S99P			
17						a173g				a224g t267a c296t E75G S99L	t314c I105T	t425g L142R	
18			a151t T51S			a172t T58S				t308c I103T		a537g a609g	
19				t168c				a205g T69Q	g209a R70K			g352c V118L	
20						a172t T58S						g465t	
21	g11a S4N					a173g				g223a E75K	t314c I105T	g475c a543g A159P	
22				t168c								t366c	t640g S214A
23			c23t g96c A8V E32D	t168a H56Q				a186t				t446c I149T	

[a]residue numbers are of hGHbp. Mutations are indicated by wild-type amino acid in single letter code followed by residue number and mutant amino acid. Nucleotide mutations are are in lower case, numbered from the first nucleotide of hGHbp. Shaded columns correspond to codons for which three or more clones had mutations.

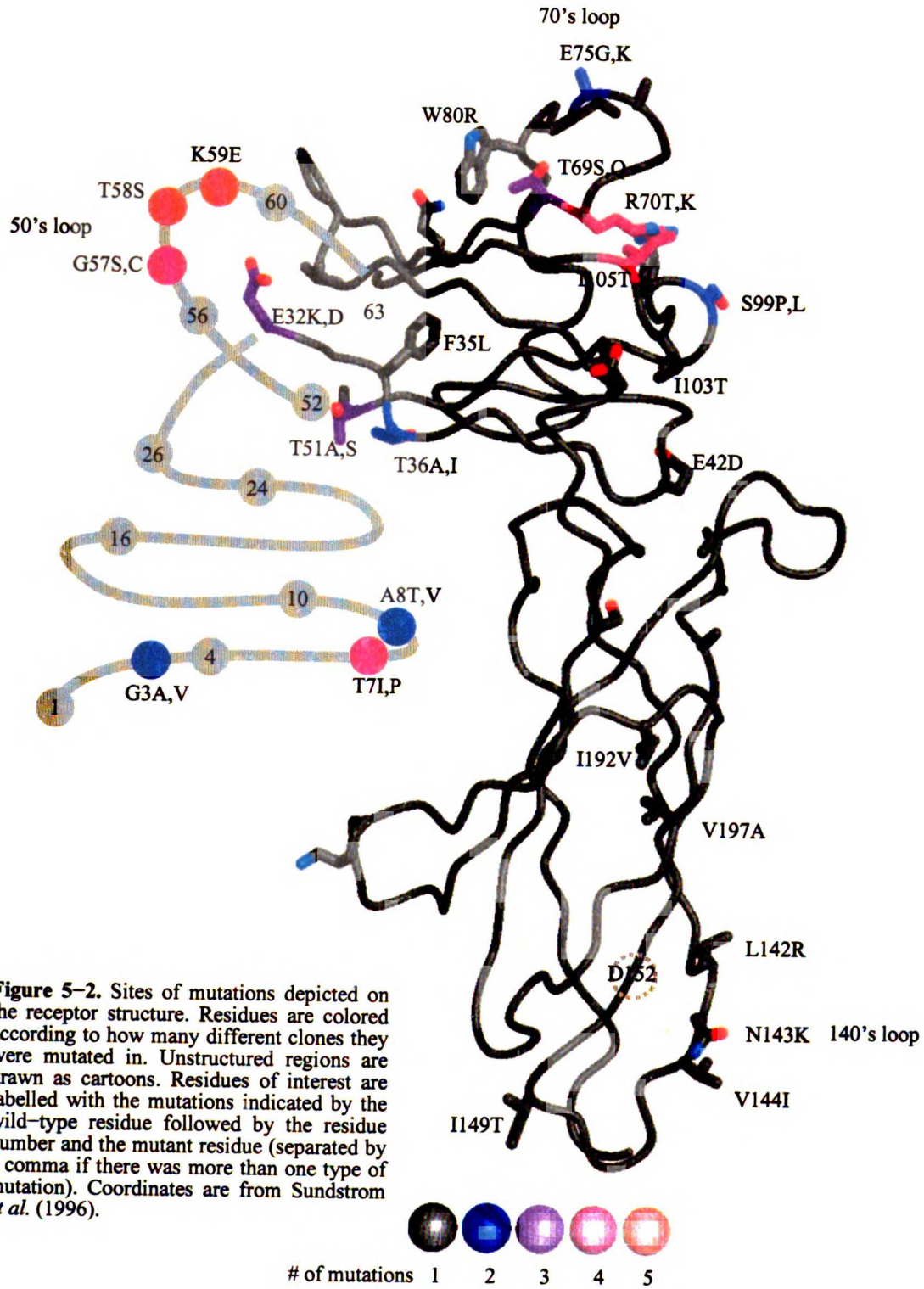

Figure 5–2. Sites of mutations depicted on the receptor structure. Residues are colored according to how many different clones they were mutated in. Unstructured regions are drawn as cartoons. Residues of interest are labelled with the mutations indicated by the wild–type residue followed by the residue number and the mutant residue (separated by a comma if there was more than one type of mutation). Coordinates are from Sundstrom *et al.* (1996).

Nine clones from round 5 and 23 clones from round 14 were fully sequenced (Table 5-2). The round 14 clones had an average of 4.1 mutations each, of which 1.6 were silent and 2.6 were missense. Most mutations were of surface residues (Figure 5-2). The ELISA signal for binding hGH showed some mutants gave a 20 to 70 fold increase in total binding activity compared to the parent hGHbp$^{1\text{-}234}$ (Table 5-3). Mutations were more concentrated in the N-terminal domain of hGHbp. The most commonly mutated residues were within or near regions of the protein that have high mobility, infered from crystallographic B-factors (Figure 5-3). This is consistent with an earlier report of DNA shuffling mutations concentrated in putative loop regions (Buchholz *et al.*, 1998). Several mutations appeared in three or more clones (referred to here as consensus mutations): T7I is in the first unstructured 32 residues of the receptor, and G57S, T58S, and K59E are in a disordered loop.

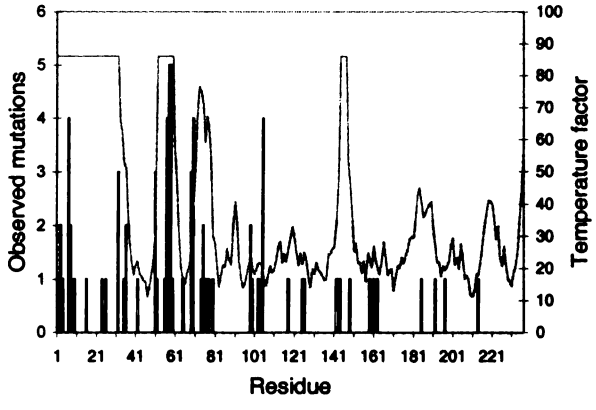

Figure 5-3. Locations of mutations versus temperature factor. The number of times a residue is mutated in libraries 5 and 14 are indicated by bars. The temperature factors of the corresponding Cα atoms in the wild-type 1:1 structure is indicated as a grey line.

The remaining consensus mutations, R70T or K and I105T, are probably involved in stabilizing the 70's loop. This loop is anchored by Arg70 at one end and Trp80 at the other end and is disordered in several receptor complexes. Ile105 is near Arg71 and in one mutant structure, rotates around to block the normal positioning of Arg71 (Chapter 2). It is possible that in the wild-type structure there is a steric clash between Ile105 and Arg71 and that the I105T and R70T, R70K mutations were selected to reduce this clash. Several other mutations might be related to the 70's loop mobility, including: T69S,Q, Q74R, E75G,K, W76R, T77S, Q78K, and W80R.

Many of the consensus mutations were to amino acids observed in homologous receptors. For example, Ile7, the first residue of the non-essential exon 3 insert (Clackson *et al.*, 1998), is more common than Thr7 among mammalian growth hormone receptors and Lys70 exists among growth hormone receptors and is exclusive in prolactin receptors (data not shown).

The information derived from several assays indicates particular roles from some of the mutations.

Several lines of evidence indicate that Mab5 binds to misfolded as well as folded hGHbp. First, an ELISA using Mab5 to capture requires 1000-fold less phage to generate a signal comparable to an ELISA using hGH to capture, despite the similar affinities of hGHbp for each (0.4 nM for Mab5 (Barnard *et al.*, 1984) and 0.3 nM for hGH) (data not shown). Second, the concentration of phage required to generate a 0.2 ELISA signal ($EC_{0.2}$) when capturing with Mab5 varies only 4-fold among the mutants, which vary 60-fold in their hGH $EC_{0.2}$. Third, the Mab5 epitope on hGHbp, inferred from these experiments and others (Duquesnoy *et al.*, 1994), appears to surround the 140's loop, which is disordered in several crystal structures and adopts alternate conformations in complexes with a single receptor bound to hGH versus complexes with two receptors bound. Fourth, Mab5 generates a 10-fold greater signal with urea solubilized inclusion bodies of histidine tagged hGHbp than with purified native hGHbp (data not shown).

Therefore we think that the ELISA signal for binding to Mab5 indicates the total displayed hGHbp, folded and unfolded. Since hGH should only binding to folded hGHbp, the ratio of hGH binding to Mab5 binding gives an indication of the percentage of hGHbp on phage that is properly folded. This ratio varied 200-fold among the mutants, with only #3 and #12 giving a ratio significantly higher than the parent. Clone #12 has two unique mutations: G3V and F35L. F35L is one of the few hydrophobic core residues to be mutated and could be responsible for the increased folding (Figure 5-2). This clone also has T7I in common with #3, the other folding-improved mutant.

Table 5-3. Properties of selectants.

clone	Increased[a] ELISA signal	Increased[b] folding	hGH IC_{50}[c]	precent[d] processed
1	9.5	0.79	4.3	20
2	1.8	0.11	-	0
3	21	2.40	1.3	20
4	3.3	0.29	2.3	0
5	5.9	*[e]	3.3	0
6	1.3	0.24	2.3	0
7	2.3	0.24	8.6	0
8	4.8	0.52	1.4	0
9	15	0.98	5.0	0[f]
10	5.3	0.70	3.3	0
11	6.2	0.49	4.2	70
12	69	19.00	2.0	0
13	5.5	0.71	1.8	10
14	9.5	1.10	7.9	20
15	7.0	0.53	1.5	20
16	2.1	0.26	9.5	10
17	3.8	0.47	6.6	10
18	1.2	0.15	21.2	0
19	4.4	0.78	3.2	50
20	6.7	0.66	1.5	10
21	24	1.40	3.2	0
22	24	1.30	1.6	20
23	37	*[e]	1.4	20

[a]indicates how much less phage is needed to generate the same ELISA signal versus $hGHbp^{1-233}$ capturing with hGH (see Materials and Methods) where the concentration in this case represents the dilution from an identically prepared phage stock.

[b]indicates the ratio of hGH $EC_{0.2}$ versus Mab5 $EC_{0.2}$ divided by the ratio for $hGHbp^{1-233}$ (see Results and Discussion). The ratio for $hGHbp^{1-234}$ is 3.1×10^{-2}.

[c]determined in a competition ELISA. The IC_{50} for $hGHbp^{1-234}$ is 1.04 nM.

[d]percentage of protein with signal peptide removed determined by western immunoblotting of total protein expressed in 33B6 cells. The parent $hGHbp^{1-233}H_6L$ had 20% processed. Clones #1,2, and 21 had much lower total expression than average.

-not measured

[e]indicates clones that gave a much lower than average maximal ELISA signal and therefore had mutations that interfered with Mab5 binding.

[f]The mobility of this protein indicated the molecular weight of a dimer.

I105T appeared in 4 clones, despite the fact that it reduced the affinity for hGH about 8-fold (Table 5-3). The beneficial effect on folding, stability, or solubility of this mutation must outweigh its negative effect on binding. I105T follows the critical Trp104 and is in a loop thought to be strained in the structure (Clackson *et al.*, 1998) and whose residues undergo significant movement in the W104A-hGHbp structure (Atwell *et al.*, 1997).

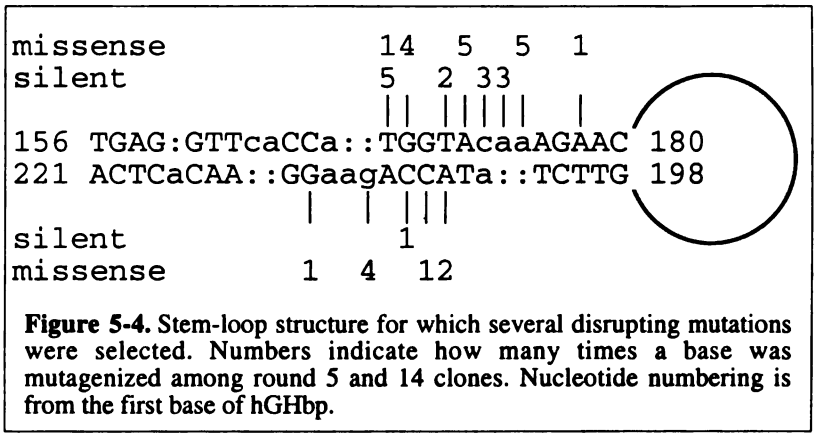

Figure 5-4. Stem-loop structure for which several disrupting mutations were selected. Numbers indicate how many times a base was mutagenized among round 5 and 14 clones. Nucleotide numbering is from the first base of hGHbp.

Two clones had mutations that significantly reduced their capture by Mab5. Clone #5 required 100-fold more phage to generate a Mab5 ELISA signal comparable to most of the clones and had a 70% reduced saturation signal. Clone #23 required 10-fold more phage and had a 30% reduced saturation signal. N143K is the only missense mutation in #5 and must be responsible for the reduced Mab5 binding. The only mutation in #23 near residue 143 is I149T. It has been observed that a mutation that confers familial growth hormone resistance, D152H, also abolishes Mab5 binding (Duquesnoy *et al.*, 1994). Therefore, Mab5 appears to have a binding epitope covering the unstable 140s loop (Figure 5-2). This loop participates in receptor-receptor interactions and the fact that Mab5 binds there explains its antagonistic activity towards hGH signaling, which occurs through dimerization of the receptors.

Two clones, #4 and #13, had only silent mutations and yet had three to five-fold improved display on phage (30 to 50-fold compared to the His tagged parent, see Table 5-3 legend). In fact, there were several silent mutations that appeared in multiple clones and were apparently selected: t168c, t171c, a173g and a174g. A secondary structure search of the gene indicated a single stable stem loop structure, encompassing these

consensus silent mutations (Figure 5-4), which was destabilized by the mutations. This stemloop could exist either in the plasmid or RNA interfering with protein expression or in the PCR reaction itself interfering with polymerization.

Since clone #11 gave a much higher level of processed protein and was therefore potentially valuable as a high expressing protein, we decided to analyze the role of its individual mutations. Each mutation was individually reverted back to wild-type. Contrary to our expectations, neither #11 nor its derivatives produced soluble protein in the periplasm under expression conditions that produced soluble wild-type hGHbp (even when the His tag was removed) (data not shown). It was found that the largest effect of the mutations was on phage yield, which is probably due to the host toxicity or phage interference of the wild-type hGHbp construct. While each mutation contributed to the overall increase in activity of the clone, only the double revertant, T10I/V192I, reduced the activity to the wild-type level, by reducing phage yield 16-fold (Table 5-4).

Table 5-4. Analysis of clone #11 by reversion mutation

mutant	phage yield[a]	increased folded display[b]	increased total display[c]	increased folding ratio[d]
hGHbp$^{1-234}$	0.09	0.5	0.3	2.1
#11[e]	1.0	1.0	1.0	1.0
" T70R[f]	0.6	0.9	1.1	0.9
" T10I	0.8	0.5	0.6	0.8
" V192I	0.5	1.0	0.8	1.2
" T10I/V192I	0.06	0.2	0.4	0.6
" t196c	0.4	1.1	1.3	0.8
" t196c/g207c	0.5	0.9	0.9	1.1

[a]quantified by OD_{268}.
[b]determined by hGH $EC_{0.2}$ using phage concentration determined by OD_{268} and compared to #11.
[c]determined by Mab5 $EC_{0.2}$ compared to #11.
[d]determined by ratio of Mab5 $EC_{0.2}$ to hGH $EC_{0.2}$ compared to the ratio for #11.
[e]all numbers were normalized to this clone.
[f]individual residues were mutated back to wild-type in the context of the rest of the mutations of #11 (see Materials and Methods).

DNA shuffling combined with sequence and functional analysis is a valuable means of obtaining functionally improved mutants and provides information about the role of particular regions and mutations in protein function. We were able to obtain mutants with greater than 60-fold improved phage display, enabling us to use phage display for engineering of hGHbp-ligand interactions. We also derived useful information about

MATERIALS AND METHODS

Reagents. Purified hGH was obtained from the Process Sciences group at Genentech.

Plasmids and oligonucleotides. The starting vector, R13, a derivative of pAMgIII (Lowman *et al.*, 1991) encoding a P_{phoA} promoter, stII signal peptide, and $hGHbp^{1-233}$-*-$gIIIp^{249-406}$ was supplied by Tim Clackson (where * refers to an amber stop codon, suppressed by Glu in XL1-Blue cells). For PCR amplification the forward primer was 5'-GGGTATCTAGAGGTTGAG-3'. For the initial library the reverse primer was 5'-AACT-GGCAAGCTTAAGACTCCTTA-3'. For subsequent rounds 5'-GCCACCGCCACTCTA-GAGGTGATGATGGTGATGATGTTTTTCGAGTGTTACATAGAG-3' was used, which introduces EKH_6L at the C-terminus of hGHbp after residue 233 for purification purposes. Derivatives of clone #11 were constructed by either site-directed or cassette mutagenesis and verified by sequencing.

DNA shuffling. Several concentrations of DNAse (1, 3, 10 µl of 0.2 mg/ml DNAse I (Sigma) in buffer (50 mM Tris (pH 7.5), 10 mM $MgSO_4$, 0.1 mM DTT, 50 µg/ml BSA, 50% glycerol) were used for limited DNAse digestion of 1 µg of the PCR amplified hGHbp gene in 100 µl 1x Vent PCR buffer (New England Biolabs) for 5 minutes. This reaction was stopped by placing in a 95 °C block. Aliquots (10 µl) were taken from each and agarose gel electrophoresis was used to determine the concentration of DNAse needed to fragment all of the full length PCR product. Following DNAse digestion, 5 µl dNTP (40 mM) and 1 µl Vent exo$^-$ polymerase (New England Biolabs) were added and PCR was performed without addition of primers (45 cycles, 30 seconds each 50 °C, 72 °C, 94 °C). Agarose gel electrophoresis was used to verify the reappearance of the full length gene product, and 10 µl was used as template to amplify in a 500 µl reaction with PCR primers and Taq polymerase (1x Taq buffer, 0.4 mM dNTP) for 15 cycles. The product was phenol/chloroform extracted, ethanol precipitated, washed and resuspended in 50 mM Tris (pH 7.5), 1 mM EDTA (TE buffer). 10 µg of the shuffled product and 10 µg of pAMgIII was digested with NsiI and XbaI overnight. The digested products were

purified by agarose gel electrophoresis and excising the correct bands. The gel bands were frozen for 30 min at -20 °C, crushed, placed in a 0.22 μm spin filter (Millipore) and centrifuged at 5,000 rpm for 20 minutes in a microcentrifuge. The gel was then extracted with 200 μl TE, and the extract recovered by centrifugation. The extracts were combined and precipitated with EtOH, washed with 70% EtOH and resuspended in 50 μl TE. After gel analysis, 0.5 μg of digested pAMgIII was ligated with 0.5 μg of the shuffled gene in ligase buffer with 1 μl T4 DNA ligase (BRL) overnight at 16° C. The ligation product was phenol/chloroform extracted, EtOH precipitated, washed and resuspended in 6 μl ultrapure water before transforming XL1-Blue cells (Stratagene) by electroporation (ref Lowman). The number of transformants for the initial, round 5 and round 10 libraries was 2.8×10^7, 1.7×10^6, and 2.2×10^6.

Phage sorting and ELISA's. Phage manipulations were carried out according to Lowman (1998). hGH was coated overnight in Maxisorp wells at 5 μg/ml in 50 mM sodium carbonate (pH 9.5), blocked for 1 hr with 2% skim milk/ PBS, and rinsed 10 times with PBS/ 0.005% tween-20 (PBST) before use in sorting or ELISA's. For sorting, hGHbp phage was incubated with immobilized hGH for 1 hour. After washing 10 to 30 times with PBST, phage was eluted with 100 μl 0.2 M glycine (pH 2), neutralized with 11 μl 1.0 M tris base and used to infect mid-log phase XL1-Blue cells. An aliquot of eluted phage was tested to determine how much phage had bound the hGH coated versus uncoated wells.

Phage ELISA's used maxisorp plates coated overnight with 5 μg/ml hGH or 1:1000 dilution of Mab5 (Barnard *et al.*, 1984). Phage serially diluted with blocking buffer was added to blocked wells, incubated for 1 hour, and rinsed with PBST before determining bound phage using the ELISA. The ELISA data from an identically prepared phage stock was fit to a binding isotherm to titer the phage. Competitive phage ELISA's utilized the amount of phage needed to give 70% of the maximal signal and competed with soluble hGH.

Proteins were expressed in *E. coli* strain 34B8, a non-suppressor host, by diluting a log-phase culture 100-fold into minimal phosphate media and growing overnight. Whole culture samples were mixed with sample buffer and run on SDS-PAGE gels under reducing conditions with purified hGHbp as a standard. Westerns were performed by transferring to PVDF membranes and using Mab5 as the primary antibody. Ratios of signal peptide processed versus unprocessed were judged visually from an autoradiograph.

Chapter 6. Summary

Phage display is an extremely powerful technique for protein engineering. The preceding chapters show its usefulness in selecting orthogonal protein binding pairs; in elucidating principles of protein function, such as the plasticity of interfaces; in selecting improved catalysts; and in elucidating display itself, through mutations in unstructured protein regions.

The future of protein engineering is bright. I believe the most important techniques in the next decade will be phage display, selection of enzymes by product capture, DNA shuffling, high-throughput screening of mutants, and computation. These techniques are compatible and the most successful engineering efforts will probably combine them. I am happy to be part of this future.

Appendix A: Protocols

Preparation of phage for assays or ssDNA sequencing. (Lowman, 1998)

1. **Inoculate 5 ml 2xYT/ 50 μg/ml carbenicillin/ 10^{10} pfu/ml VCS helper phage in a 15 ml tube with a single, well-separated colony.** *It is best to use freshly grown colonies.*

2. **Grow overnight on a ferris wheel at 37° C.**

3. **Spin down cells 10k rpm, 10 minutes. Prepare conical tubes with 1.5 ml 20% PEG-8000 / 2.5 M NaCl (PEG/NaCl).**

4. **Pour supernatent into PEG/NaCl tubes and vortex to mix.** *Optional: leave at room temperature for 15 minutes.*

5. **Spin precipitated phage 5k rpm, 10 minutes, decant, respin, and aspirate remaining supernatant with vacuum line and microtip.**

1. **Resuspend pellet in 200 μl PBS and transfer to an epindorf.** *Can use 100 to 1000 μl of PBS depending on the phage pellet size and desired phage concentration.*

7. **Use directly for phage ELISA.** *Optional: respin to remove debris, transfer supernatant, repeat phage precipitation by adding 1/4 volume PEG/NaCl, spin, and resuspend in PBS.*

8. **For sequencing ssDNA, heat at 100 °C for 10 min., spin in microfuge to remove denatured protein, and transfer supernatent to fresh epindorf tube for ssDNA sequencing.**

Large scale Kunkel mutagenesis. (Sidhu & Wells, in preparation)

1. **Incubate oligo with kinase at 37 °C for 1 hour.**

Rxn: 2 µl oligo (330 ng/µl stock, $A_{260} \cong 10$)
2 µl 10x TM buffer (0.5 M tris (pH 7.5), 0.1 M $MgCl_2$)
2 µl 10 mM ATP
1 µl 100 mM DTT
12.5 µl H_20
0.5 µl kinase (New England Biolabs, 10 U/µl)

2. **Anneal oligo with template by heating at 85 °C for 30 seconds and then slowly cooling to room temperature in 20 minutes.** *This can be accomplished by preheating a water bath on a hot plate to 85°C, incubating the annealing reaction 30 seconds , and removing the water bath to the bench top. Due to dephosphorylation, the oligo should be used soon after the kinase reaction and should not be heated for a prolonged period.*

Rxn: 20 µg kunkel template (1 x 10^{13} molecules for a 5 kb template)
0.6 µg kinased oligo (20 µl of the above reaction; oligo/template ≅ 3)
25 µl 10x TM buffer
H_20 to 250 µl final volume

3. **Fill in the remaining nucleotides by adding polymerase and nucleotides and incubating 20 °C for 3 hours.**

Add: 1 µl 100 mM ATP
10 µl 25 mM dNTPs (25 mM each dATP, dCTP, dGTP, dTTP)
15 µl 100 mM DTT
6 µl T4 ligase (NEB, 400 Uµl)
3 µl T7 polymerase (NEB, 10 Uµl)

4. **Run 1 µl on a 0.7 to 1% agarose gel buffered with TAE.** *The products are not as distinct when using TBE. Three bands are usually visible. The fastest migrating band is correctly filled and ligated and after transformation yields mostly mutants. The middle band is filled but unligated and yields mostly wild-type tranformants. The slowest band is strand displaced and transforms 30-fold less efficiently than the others.*

5. **Extract fill-in reaction with an equal volume of phenol/chloroform and then an equal volume of chloroform.**

6. **Purify DNA using QIAquick gel extraction kit.** *Use one column for each 10 µg of template DNA. Elute DNA from column with 30 µl H_20. Optional: analyze 0.5 µl on an agarose gel.*

Large scale electroporation. (Sidhu & Wells, in preparation)

1. **Thaw 2x 350 μl electrocompetent *E. coli* SS320 cells on ice.**

2. **Mix 30 μl of DNA with 350 μl cells, inject into chilled 0.2 cm gap cuvette and electroporate using a Biorad electroporater. Settings: 2.5 kV, 200 ohms, 25 μF.**

3. **Quickly transfer to 25 ml of SOC and incubate with shaking at 37°C for 1 hour.**

4. **Remove 100 μl for titrating transformed cells.**

5. **Transfer culture to 500 ml 2xYT / 50 μg/ml carbenicillin/ 10^{10} pfu/ml VCS helper phage. (Grow overnight and harvest phage from supernatant using a 1/4th volume of PEG/NaCl to precipitate.)**

6. **Finish titrating transformed cells either using serial 100-fold dilutions with 2x YT and plating 100 μl on separate plates, or using serial 10-fold dilutions in a microtiter wells and plating a series of 10 μl spots.** *It is important to change tips after every dilution and to mix dilutions thoroughly. If care is taken to minimize contaminants in preparing cells and DNA, $2x10^{10}$ transformants are typical.*

Preparation of electrocompetent *E. coli.* (Sidhu & Wells, in preparation)

1. **Inoculate 1 ml 2xYT/ 5 μg/ml tetracycline with a single fresh colony of SS-320 (*E. coli* MC1061 with the F-factor from XL1-Blue (Stratagene)) grown on LB/ 5 μg/ml tetracycline. Grow about 6 hours and inoculate 50 ml 2x YT/ 5 μg/ml tetracycline in a 500 ml flask. Grow overnight.**

2. **Inoculate 6 x 900 ml superbroth/ 5 μg/ml tetracycline in 2 L baffled flasks with 5 ml of overnight culture. Grow to $OD_{600} \cong 0.7$ at 37C, 200 rpm shaking (approximately 3 to 4 hours).**

3. **Chill three flasks 10 minutes on ice, shaking periodically.**

All subsequent steps are performed in a cold room, on ice with prechilled solutions and equipment.

4. **Centrifuge in 400 ml centrifuge tubes 5.5k rpm 5 minutes (start chilling remaining flasks) and decant. Add culture from remaining three flasks to the same centrifuge tubes. Respin and decant.**

5. **Resuspend in an equal volume of 1 mM HEPES (pH 7.4).** *Use "sterile water for irrigation USP". Use stir bars soaked in ethanol. Swirl first with stir bar to break up pellet and then place on stir plate to homogenize with a moderate stir rate until pellet is completely resuspended.*

6. **Centrifuge 5.5k rpm 10 minutes, leaving stir bar in tube.** *Be careful not to disturb the pellet when removing the tube from the rotor.*

7. **Resuspend in an equal volume of 1 mM HEPES as above and respin.**

8. **Resuspend each pellet in 100 ml 10% glycerol.** *Do not combine cells. Use irrigation water and ultrapure glycerol (Gibco) and filter sterilize.*

9. **Centrifuge 5.5k rpm 15 minutes and carefully decant all supernatant, removing traces with a pipet. Remove stir bar.**

10. **Resuspend in a minimal volume (3 mls) of 10% glycerol.**

11. **Aliquote 350 μl into eppendorf tubes and freeze in a dry ice/ ethanol bath. Store at –70°C.**

***B. subtilis* transformation.** (Anagnostopoulos & Spizizen, 1961)

1. **Restreak *B. subtilis* BG2036 on LB agar plates (no antibiotics). Grow overnight at 37°C.**

2. **Inoculate 2.5 ml minimal glucose (MG) media supplemented with 25 µl of 5 mg/ml tryptophan and 2.5 µl of 20% casamino acids. Grow with good aeration 37°C 4 to 5 hours until very turbid.** *B. subtilis* reaches competence in stationary phase.

3. **Dilute 10-fold into 5 ml MG supplemented with 2.5 µl tryptophan and 1.3 µl casamino acids.**

4. **Add approximately 1 µg dsDNA to 300 µl cells.** DNA should be prepared from strains that produce multimeric plasmids. JM101, MM294, and GM48 work well.

5. **Grow with aeration 1 to 2 hours at 37 °C.**

6. **Plate on selective media and grow overnight.** For pSS5 based plasmids, chloramphenicol (10 µg/ml) or kanamycin.

Bacillus Minimal Glucose:

Per 1 L: 14 g K_2HPO_4, 6 g KH_2PO_4, 2 g $(NH_4)_2SO_4$, 1 g $Na_2Citrate.2H_2O$, distilled H_2O to 1 L, dissolve, autoclave, cool to 55°C

Add: 2.5 ml 1 M $MgSO_4$, 2.0 ml 2.5 mM $FeCl_3$, 2.0 ml 10 mM $MnCl_2$, 10 ml 50% glucose

Subtilisin and subtiligase preparation. (Carter & Wells, 1988)

1. **Streak expression strain on LB/ 10 μg/ml chloramphenicol. Grow overnight at 37°C.**

2. **Inoculate 25 ml 2xYT/ 10 μg/ml chloramphenicol with a single colony. Grow at 37°C with shaking to approximately 1 OD_{600}.**

3. **Inoculate 5 ml of culture into 500 ml 2xYT/ chloramphenicol. Grow at 37°C with shaking for 24 hours.**

All subsequent steps are performed on ice or in a cold room.

4. **Centrifuge cells at 4°C 8k rpm for 10 minutes.** *To avoid cell lysis it is important that cultures are spun immediately after removing them from shaking.*

5. **Transfer supernatant to a beaker and add an equal volume of cold ethanol. Mix thoroughly.**

6. **Centrifuge at 8k rpm for 10 minutes.**

7. **Transfer supernatant to a beaker and add an equal volume of cold ethanol (to 75% ethanol). Mix thoroughly.**

8. **Centrifuge 5k rpm 5 minutes.** *The lower speed makes it easier to resuspend the crude subtilisin pellet.*

9. **Decant and resuspend pellet in 10 ml of 50 mM Tris (pH 8.0)/ 5 mM $CaCl_2$/ 10 mM DTT** (to reduce the active site cysteine in subtiligase).

10. **Clarify at 40,000g for 30 minutes.**

11. **Transfer the supernatant to a new centrifuge tube and add ammonium sulfate to 70% saturated to precipitate subtilisin.** Use 3.5 volumes of 90% saturated ammonium sulfate.

12. **Resuspend in 5 ml buffer for final column purification.**

13. **Purify subtiligase by resuspending in NaAcetate (pH 4.8)/ 5 mM $CaCl_2$/ 100 mM NaCl/ 1 mM PMSF.** *Purification is performed at room temperature. The low pH stabilizes the disulfide between subtiligase and the column.*

14. **Load onto 2 ml of pre-equilibrated thiol phenyl sepharose 6B (Pharmacia).**

15. **Rinse with 5 ml buffer. Bump with 5 ml buffer + 5 mM DTT. Rinse with 5 ml buffer.** *The second buffer rinse removes DTT before raising the pH and reduces unwanted elution.*

16. **Raise pH with 5 ml 50 mM Tris (pH 7.5)/ 5 mM $CaCl_2$/ 100 mM NaCl.** *At pH 7.5 subtiligase is bound tighter and elutes later than at pH 8.0.*

17. **Elute with step gradient of 2.5 ml buffer + 5, 10, 20 mM DTT, followed by several 2.5 ml fractions of buffer + 20 mM DTT.** *The elutes in the second 20 mM DTT fraction for wild-type subtiligase.*

18. **Use PD10 columns to exchange the buffer to 10 mM Tricine (pH 8.0)/ 5 mM $CaCl_2$.**

19. **Protein stocks are flash frozen in a dry ice/ ethanol bath and stored at –70 °C.** *Subtiligase should be used in a reducing buffer to avoid active site oxidation. 1 mM DTT does not interfere with ligation.*

Cassette subcloning

1. **Mix the PCR reaction in a PCR tube.**

Rxn: 42 µl H_2O (distilled water stored in plastic)
5 µl 10x PCR buffer
1 µl 10 mM dNTP (2.5 mM each)
1 µl each N- and C-terminal oligonucleotide primers (A_{260} = 10 stock)

2. **Put the reaction in a thermocyler. Heat the mix 94 °C 3 minutes. Add 0.5 µl Taq polymerase after 2 minutes.**

3. **Cycle the reaction: 25x 94 °C 1 minute, 55 °C 1 minute, 72 °C 1 minute/1000 bp.**

4. **Purify the PCR product using Qiagen PCR purification kit.**

5. **Double digest the purified PCR product and the vector (1 µg) with 1 µl each of the endonucleases in the appropriate buffer overnight (20 µl final).**

6. **Add 1 µl of shrimp alkaline phosphatase to the vector digestion and incubate at 37 °C for 1 hour.** *This reduces background.*

7. **Heat kill the phosphotase at 65 °C for 15 minutes.**

8. **Run digested vector and PCR product on an agarose gel. Purify the appropriate band using a Qiagen gel purification kit** (or flash freeze in an EtOH/dry ice bath for five minutes, thaw at 37°C for 5 minutes, crush and spin liquid through a millipore filter).

9. **Mix insert and vector to give a 3:1 molar ratio with 2 µl 5x ligase buffer (9 µl final) and add 1 µl T4 DNA ligase. Incubate at room temperature for 2 hours.**

10. **Transform PEG/DMSO competent cells (20 µl) with 2 µl of reaction in 20 µl 1x KCM.**

PEG/DMSO Trasformation (Chung & Miller, 1988)

Preparation of competent cells:

1. **Grow cells in LB to OD_{600} = 0.3 to 0.6.** From an overnight culture, dilute saturated cells 100-fold in LB and grow at 37°C with 250 rpm shaking for about 2 hours.

2. **Centrifuge 3k rpm 5 minutes 4°C.**

3. **Resuspend in 1/20 volume transformation and storage buffer (TSB).**

TSB: LB (pH 6.1 with HCl)
10% PEG 3,350
5% DMSO
10 mM $MgCl_2$
10 mM $MgSO_4$

4. **Incubate cells on ice for 10 minutes.**

5. **Add glycerol to 10%, aliquot 50 ml into eppendorf tubes and freeze in a dry ice/ ethanol bath. Store at –70 °C.**

Transformation:

1. **Dilute DNA 10-fold with KCM buffer.**

KCM: 0.1 M KCl
30 mM $CaCl_2$
50 mM $MgCl_2$

2. **Mix DNA/KCM with an equal volume of thawed cells. Place on ice for 20 minutes.**

3. **Incubate at room temperature for 10 minutes.**

4. **Add 300 µl of LB and incubate at 37 °C for 1 hr.** *This incubation period is necessary for attaining resistance to cytotoxic antibiotics and for resolving mutliply transformed bacteria, especially when using kunkel mutagenesis.*

5. **Plate on appropriate antibiotic plate and incubate at 37 °C overnight.**

Preparation of ssDNA for Kunkel mutagenesis

1. **Transform a dut$^-$ ung$^-$ strain (CJ236) with the target plasmid. Plate on LB with 10 μg/ml chloramphenicol for maintaining the F factor in CJ236 and the appropriate antibiotic (50 mg/ml carbanecilin) for the plasmid.**

2. **Inoculate 25 to 500 ml 2xYT/ 50 μg/ml carbenicillen/ 10^{10} pfu/ml VCS helper phage. Grow overnight.**

3. **Centrifuge at 8k rpm for 10 minutes to remove cells.**

4. **Precipitate phage with 1/4 volume of PEG/NaCl. Centrifuge 8k rpm for 10 minutes.**

5. **Resuspend phage pellet in PBS and purify ssDNA using Qiagen m13 ssDNA purification kit.**

Preparation of helper phage stock

1. **Grow XL1-Blue cells in 2xYT/ 10 μg/ml tetracycline to OD_{600} = 0.8.**
2. **Dilute a stock of helper phage (M13KO7 or VCS) 1000-fold with 2xYT.**
3. **Streak out the diluted phage on an LB agar plate with a sterile loop using zigzags in one direction.**
4. **Mix 500 μl XL1-Blue culture with 4 ml of melted 2xYT top agar and quickly pour the mixture across the plate starting from the opposite side of the plate as the phage streak. Incubate at 37°C overnight.**
5. **Pick a single well-separated plaque with a toothpick and place into 1 ml of 2xYT with 50 mg/ml kanamycin and 30 mg/ml tetracycline. Incubate 12-16 hrs at 37 °C and 250 rpm shaking.**
6. **Centrifuge 5k rpm 10 minutes and combine supernatant with 1 ml of fresh XL1-Blue culture (OD_{600} ~ 0.8). Incubate 37 °C on a ferris wheel for 30 minutes.**
7. **Transfer mixture to a 250 ml flask with 25 ml 2xYT and 50 μg/ml kanamycin. Grow at 37°C with shaking for 20-24 hours.**
8. **Centrifuge 10k rpm 10 minutes and transfer supernatant to a fresh centrifuge tube. Add 1/4 volume of PEG/ NaCl. Let stand for 10 minutes.**
9. **Centrifuge 10k rpm 10 minutes, decant supernatant, respin for 2 minutes, aspirate supernatant and dissolve phage pellet in 1 ml PBS. Spin for 5 minutes to remove undissolved material. Add 200 μl of 20% PEG/ 2.5 M NaCl, sit for 10 minutes, aspirate, respin, aspirate.**
10. **Dissolve phage pellet in 25 μl PBS. Store at 4 °C.**

Determining phage titer:

1. **Grow XL1-Blue cells in 2xYT / 10 μg/ml tetracycline to OD_{600} = 0.8**
2. **Thoroughly melt 2xYT top agar in the microwave and then equilibrate the melted agar to 48-50° C.** *Loosen the cap on the bottle before heating.*
3. **Dilute original phage stock by serial dilutions (20 μl to 180 μl PBS to 10^{12}-fold dilution).** *Use a new pipet tip for each dilution.*
4. **In miniprep tubes, combine 200 μl XL-1 cells with 100 μl duluted phage (usually 10^9).**
5. **Add 4 ml of 48-50° C melted agar to tubes and quickly pour to warmed LB agar plates.**
6. **Incubate at 37°C overnight and count plaques in the morning.**

Phage ELISA. (Lowman, 1998)

1. **Coat microtiter plate wells with 100 µl of 5 µg/ml target protein at 4 °C overnight.**

2. **Discard coat protein solution and block wells with 400 µl Casein Blocking buffer (Pierce) or 6% skim milk/PBS for 1 hour at room temperature.**

3. **Rinse 10x with PBS/0.005%tween-20 (PBST).**

4. **In an non-stick plate prepare 4-fold serial dilutions of phage in blocking buffer (50 µl phage + 150 µl buffer).**

5. **Transfer 100 µl of phage titration to coated and blocked plate and incubate 1-2 hours.**

6. **Rinse 10x with PBST and add 1/5000 dilution of α-m13 HRP (???). Incubate for 20 minutes.**

7. **Rinse 10x with PBST. Rinse once with PBS. Add OPD substrate (Sigma) (1 pill / 4 µl H_2O_2 / 10 ml PBS).**

8. **Develop 10 minutes or until colored with shaking. Add 1/2 volume 2.5 M H_2SO_4 to stop reaction.**

9. **Read plate at 492 nm.**

10. **For a competitive ELISA use the least amount of phage needed to give a trustworthy signal (e.g. OD_{492} = 0.1), and titrate competitor in phage ELISA.** *At low phage concentrations the observed IC_{50} approaches Kd + [target] (DeLano, 1999).*

PCR Shuffling

1. **Use a large scale PCR reaction to amplify gene of interest. Use oligos with >5 bp beyond the restriction sites used for subcloning. Run analytical agarose gel.**

2. **Wizard PCR purify product.**

3. **DNAse treat PCR product. In a PCR microcentrifuge tube mix 10 μl PCR product, 10 μl vent exo⁻ buffer, 75 ml H20. Preheat thermocycler to 94 C.**

4. **Add 5 μl +/- 4 μl DNAse (0.2 mg/ml [0.1 unit/ml] in 50 mM Tris (7.5), 10 mM $MgSO_4$, 0.1 mM DTT, 50 mg/ml BSA, 50% glycerol, stored at -20 °C).**

5. **Treat DNA with DNAse for 5 minutes at room temperature and transfer to 94 °C thermocycler. Run analytical agarose gel on all DNAse concentrations.**

6. **Add 5 μl dNTP, 1 μl Vent exo- polymerase to all reactions.** *Vent makes 70% blunt product, important for assembly intermediates, and exo- gives a higher error rate.*

7. **Perform PCR reaction: 45x (94 °C 30 sec, 55 °C 30 sec, 72 °C 30 sec).**

8. **Use reaction derived from most thoroughly digested DNA to amplify with standard PCR reaction, using Taq polymerase and primers. Only 15 cycles.**

9. **Phenol/chloroform extract product. Precipitate with 3.5 volumes cold EtOH, wash with 70% EtOH and resuspend in 30 μl TE.**

10. **Digest 16 μl with 1 μl each endonuclease for subcloning in appropriate buffer overnight. Digest vector if necessary.**

11. **Gel purify product.**

12. **Ligate shuffled gene with vector.**

    10 μl cut vector (1 μg)
    9 μl cut insert (0.5 μg)
    5 μl ligation buffer (BRL)
    on ice
    Add 1 μl ligase (BRL)
    16° C overnight

13. **Transform *E. coli*.**

Bibliography

Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakoff, A. & Wells, J. A. (1991) Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**, 4151-4159.

Anagnostopoulos, C. & Spizizen, J. (1961) Requirements for transformation in *Bacillus subtilis J. Bact.* **81**, 741-746.

Aota, S.-i., Gojobori, T., Ishibashi, F., Maruyama, T., & Ikemura, T. (1988). Codon usage tabulated from the GenBank genetic sequence database. *Nucleic Acids Res. suppl.* **16**, 315–402.

Arkin, M. R. & Wells, J. A. (1998) Probing the importance of second sphere residues in an esterolytic antibody by phage display.*J. Mol. Biol.* **284**, 1083-1094.

Atwell, S., Ridgway, J. B. B., Wells, J. A. & Carter, P. (1997) Stable heterodimers from remodeling the domain interface of a homodimer using a phage display library. *J. Mol. Biol.* **270**, 26-35.

Atwell, S., Ultsch, M., de Vos, A. M. & Wells, J. A. (1997) Structural plasticity in a remodeled protein-protein interface. *Science* **278**, 1125-1128.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (eds.) (1996). In *Current Protocols in Molecular Biology.* John Wiley, New York, NY.

Baca, M., Scanlan, T. S., Stephenson, R. C. & Wells, J. A. (1997) Phage display of a catalytic antibody to optimize affinity for transition-state analog binding. *Proc. Natl. Acad. Sci, USA* **94**, 10063-10068.

Barany, G. & Merrifield, R. B. (1980) in *The Peptides*, eds. Gross, E. & Meinenhoffer, J. (Academic, New York), Vol. 2, pp. 1-284.

Barnard, R., Peter, G. B., Rylatt, D. B. & Waters, M. J. (1984) Monoclonal antibodies to

the rabbit liver growth hormone receptor: production and characterization. *Endocrinol.* **115**, 1805-1813.

Bass, S. H., Mulkerrin, M. G. & Wells, J. A. (1991) A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor. *Proc. Natl. Acad. Sci. USA* **88**, 4498-4502.

Bass, S., Greene, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins: Struct. Funct. Genet.* **8**, 309–314.

Bedouelle, H. & Winter, G. (1986). A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature* **320**, 371–373.

Belshaw, P. J., Schoepfer, J. G., Liu, K-Q., Morrison, K. L. & Schreiber, S. L. (1995). Rational design of orthogonal receptor-ligand combinations. *Angew. Chem. Int. Ed. Engl.* **34**, 2129–2132.

Bender, M. L., Clement, G. E., Gunter, C. R. & Kezdy, F. J. (1964) The kinetics of α-chymotrypsin reactions in the presence of added nucleophiles. *J. Am. Chem. Soc.* **86**, 3697-3703.

Berman, P. W., Gregory, T. Crase, D. & Lasky, L. A. (1985). Protection from genital herpes simplex virus type 2 infection by vaccination with cloned type 1 glycoprotein D. *Science* **227**, 1490–1492.

Bodanszky, M. & Bodanszky, A. (1994) in *The Practice of Peptide Synthesis, 2nd Ed. rev.* (Springer-Verlag, New York).

Bongers, J. & Heimer, E. P. (1994) Recent applications of enzymatic peptide synthesis. *Peptides* **15**, 183-193.

Bott, R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B. & Power, S. (1988) The three-dimensional structure of *Bacillus amyloliquefaciens* subtilisin at 1.8 Å and an analysis of the structural consequences of peroxide inactivation. *J. Biol. Chem.* **263**, 7895-7906.

Bourell, J. H., Clauser, K. P., Kelley, R., Carter, P. & Stults, J. T. (1994) Electrospray ionization mass spectrometry of recombinantly engineered antibody fragments. *Anal. Chem.* **66**, 2088–2095.

Boyd, D. & Beckwith, J. (1990) The role of charged amino acids in the localization of secreted and membrane proteins. *Cell* **62**, 1031-1036.

Braisted, A. C., Judice, J. K. & Wells, J. A. (1997) Synthesis of proteins by subtiligase. *Methods Enzymol.* **289**, 298-313.

Bratovanova, E. K. & Petkov, D. D. (1987) Glycine flanked by hydrophobic bulky amino acid residues as minimal sequence for effective subtilisin catalysis. *Biochem. J.* **248**, 957-960.

Braxton, S. & Wells, J. A. (1991) The importance of a distal hydrogen bonding group in stabilizing the transition state in subtilisin BPN'. *J. Biol. Chem.* **266**, 11797-11800.

Bromme, D., Peters, K., Fink, S. & Fittkau, S. (1986) Enzyme-substrate interactions in the hydrolysis of peptide substrates by thermitase, subtilisin BPN', and proteinase K. *Arch. Biochem. Biophys.* **244**, 439-446.

Brunger, A. T., Kuriyan, J. & Karplus, M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458-460.

Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G. & Gallagher, D. T. (1992) Energetics of folding subtilisin BPN'. *Biochemistry* **31**, 4937-4945.

Buchholz, F., Angrand, P.-O. & Stewart, A. F. (1998) Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nature Biotech.* **16**, 657-662.

Byrn, R. A., Mordenti, J., Lucas, C., Smith, D., Marsters, S. A., Johnson, J. S., Cossum, P., Chamow, S. M., Wurm, F. M., Gregory, T., Groopman, J. E. & Capon, D. J. (1990). Biological properties of a CD4 immunoadhesin. *Nature* **344**, 667–670.

Cadwell, R. C. & Joyce, G. F. (1992) Randomization of genes by PCR mutagenesis. *PCR Meth. Appl.* **2**, 28-33.

Carcamo, J., Ravera, M. W., Brissette, R., Dedova, O., Beasley, J. R., Alam-Moghe, A., Wan, C., Blume, A. & Mandecki, W. (1998) Unexpected frameshifts from gene to expressed protein in a phage-displayed peptide library. *Proc. Natl. Acad. Sci. USA* **95**, 11146-11151.

Carter, P., Bedouelle, H. & Winter, G. (1986). Construction of heterodimer tyrosyl-tRNA synthetase shows $tRNA^{Tyr}$ interacts with both subunits. *Proc. Natl. Acad. Sci. USA* **83**, 1189–1192.

Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D., Wong, W. L. T., Rowland, A. M., Kotts, C. E., Carver, M. E., Yang, M., Bourell, J. H., Shepard, H. M. & Henner, D. (1992). High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. *Bio/Technology*, **10**, 163–167.

Carter, P., Nilsson, B., Burnier, J. P., Burdick, D. & Wells, J. A. (1989). Engineering subtilisin BPN_ for site-specific proteolysis. *Proteins: Struct. Funct. Genet.* **6**, 240–248.

Carter, P., Ridgway, J. & Zhu, Z. (1995). Toward the production of bispecific antibody fragments for clinical applications. *J. Hematotherapy* **4**, 463–470.

Carter, P. & Wells, J. A. (1987) Engineering enzyme specificity by "substrate-assisted catalysis". *Science* **237**, 394-399.

Chamow, S. M., Zhang, D. Z., Tan, X. Y., Mhatre, S. M., Marsters, S. A., Peers, D. H., Byrn, R. A., Ashkenazi, A. & Junghans, R. P. (1994). A humanized, bispecific immunoadhesin-antibody that retargets $CD3^+$ effectors to kill HIV-1-infected cells. *J. Immunol.* **153**, 4268–4280.

Chang, C. N., Kuang, W.-J. & Chen, E. Y. (1986). Nucleotide sequence of the alkaline phosphatase gene of *Escherichia coli. Gene* **44**, 121–125.

Chang, C. N., Rey, M., Bochner, B., Heyneker, H. & Gray, G. (1987) High-level secretion of human growth hormone by Escherichia coli. *Gene* **55**, 189-196.

Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. (1994) Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA* **91**, 12544-12548.

Chung, C. T. & Miller, R. H. (1988) A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res.* **16**, 3580.

Clackson, T. & Wells, J. A. (1994) In vitro selection from protein and peptide libraries. *Trends Biotechnol.* **12**, 173-184.

Clackson, T. & Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383-386.

Clackson, T., Ultsch, M. H., Wells, J. A. & de Vos, A. M. (1998) Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity. *J. Mol. Biol.* **277**, 1111-1128.

Collaborative Computational Project, Number 4, 1994 (1994) The CCP4 suite: programs for protein crystallography. *Acta Cryst.* **D50**, 760-763.

Corey, D. R., Shiau, A. K., Yang, Q., Janowski, B. A. & Craik, C. S. (1993) Trypsin display on the surface of bacteriophage. *Gene* **128**, 129-134.

Creighton, T. E. (1993). In *Proteins: Structure and Molecular Properties,* 2nd ed., Freeman, New York, NY.

Crissman, J. W. & Smith, G. P. (1984) Gene-III protein of filamentous phages: evidence for a caboxyl-terminal domain with a role in morphogenesis. *Virrology* **132**, 445-455.

Cunningham, B. C. & Wells, J. A. (1989) High-resolution epitope mapping of hGH-receptor interactions by ananine-scanning mutagenesis. *Science* **244**, 1081-1085.

Cunningham, B. C. & Wells, J. A. (1993) Comparison of a structural and a functional epitope. *J. Mol. Biol.* **234**, 554-563.

Cunningham, B. C., Jhurani, P., Ng, P. & Wells, J. A. (1989) Receptor and antibody epitopes in human growth hormone identified by homolog-scanning mutagenesis. *Science* **243**, 1330-1336.

Cunningham, B. C., Lowe, D. G., Li, B., Bennett, B. D. & Wells, J. A. (1994) Production of an atrial natriuretic peptide variant that is specific for type A receptor. *EMBO J.* **13**, 2508-2515.

Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779.

de Vos, A. M., Ultsch, M. & Kossiakoff, A. A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306-312.

Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9- and 2.8 Å resolution. *Biochemistry* **20**, 2361–2370.

DeLano, W. (in preparation) Ph. D. University of California, San Francisco.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L, Neri, P., Winter, G. & Neri, D. (1999) A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage. *J. Mol. Biol.* **286**, 617-633.

Deonarain, M. P., Scrutton, N. S. & Perham, R. N. (1992). Engineering surface charge. 2. A method for purifying heterodimers of *Escherichia coli* gluathione reductase. *Biochemistry* **31**, 1498–1504.

Dimasi, N., Pasquo, A., Martin, F., Di Marco, S., Steinkuhler, C., Cortese, R. & Sollazzo, M. (1998) Engineering, characterization and phage display of hepatitis C virus NS3 protease and NS4A cofactor peptide as a single-chain protein. *Prot. Eng.* **11**, 1257-1265.

Duquesnoy, P., Sobrier, M.-L., Duriez, B., Dastot, F., Buchanan, C. R., Savage, M. O., Preece, M. A., Craescu, C. T., Blouquit, Y., Goossens, M & Amselem, S. (1994) A single amino acid substitution in the exoplasmic domain of the human growth hormone (GH) receptor confers familial GH resistance (Laron syndrome) with

positive GH-binding activity by abolishing receptor homodimerization. *EMBO J.* **13**, 1386-1395.

Ellerson, J. R., Yasmeen, D., Painter, R. H. & Dorrington, K. J. (1976). Structure and function of immunoglobulin domains. III. Isolation and characterization of a fragment corresponding to the C_2 homology region of human immunoglobin G1. *J. Immunol.* **116**, 510–517.

Ellison, J. W., Berson, B. J. & Hood, L. E. (1982). The nucleotide sequence of a human immunoglobulin $C_{\gamma 1}$ gene. *Nucleic Acids Res.* **10**, 4071–4079.

Elsevier, J. P., Wells, L., Quimby, B. B & Fridovich-Keil, J. L. (1996). Heterodimer formation and activity in the human enzyme galactose-1-phosphate uridylyltransferase. *Proc. Natl. Acad. Sci. USA* **93**, 7166–7171.

Eriksson, A. E., Baase, W. A. & Matthews, B. W. (1993) Similar hydrophobic replacements of Leu99 and Phe153 within the core of T4 lysozyme have different structural and thermodynamic consequences. *J. Mol. Biol.* **229**, 747-769.

Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **255**, 178-183.

Estell, D. A., Graycar, T. P. & Wells, J. A. (1985) Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.* **260**, 6518-6521.

Fersht, A. (1985) *Enzyme Structure and Mechanism* (W. H. Freeman & Company, New York).

Figini, M., Marks, J. D., Winter, G. & Griffiths, A. D. (1994). *In vitro* assembly of repertoires of antibody chains on the surface of phage by renaturaion. *J. Mol. Biol.* **239**, 68–78.

Fujii, I., Fukuyama, S., Iwabuchi, Y. & Tanimura, R. (1998) Evolving catalytic antibodies in a phage-displayed combinatorial library. *Nat. Biotechnol.* **16**, 463-467

Gallagher, T., Gilliland, G., Wang, L. & Bryan, P. (1995) The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* **3**, 907-914.

Gao, C., Lin, C.-H., Lo, C.-H. L., Mao, S., Wirsching, P., Lerner, R. A. & Janda, K. D. (1997) Making chemistry selectable by linking it to infectivity. *Proc. Natl. Acad. Sci. USA* **94**, 11777-11782.

Garrard, L. J., Yang, M., O'Connell, M. P., Kelley, R. F. & Henner, D. J. F_{AB} assembly and enrichment in a monovalent phage display system. (1991). *Bio/Technology*, **9**, 1373–1377.

Griffiths, A. D. & Duncan, A. R. (1998) Strategies for selection of antibodies by phage display. *Curr. Opin. Biotechnol.* **9**, 102-108.

Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L., Kontermann, R., Jones, P. T., Low, N. M., Allison, T. J., Prospero, T. D., Hoogenboom, H. R., Nissim, A., Cox, J. P. L., Harrison, J. L., Zaccolo, M., Gherardi, E. & Winter, G. Isolation of high affinity human antibodies directly from large synthetic repertoires. (1994). *EMBO J.* **13**, 3245–3260.

Gron, H., Meldal, M. & Breddam, K. (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching. *Biochemistry* **31**, 6011-6018.

Gron, H., Meldal, M. & Breddam, K. (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching. *Biochemistry* **31**, 6011-6018.

Hansson, L. O., Widersten, M. & Mannervik, B. (1997) Mechanism-based phage display selection of active-site mutants of human glutathione transferase A1-1 catalyzing SNAr reactions. *Biochemistry* **36**, 11252-11260.

Holford, M. & Muir, T. W. (1998) Adding 'splice' to protein engineering. *Structure*, **6**, 961-956.

Jacobsson, K. & Frykberk, L. (1995) Cloning of ligand-binding domains of bacterial receptors by phage display. *Biotechniques* **18**, 878-885.

Janda, K. D., Lo, C., Li, T., Barbas, C. F., Wirsching, P. & Lerner, R. A. (1994) Direct selection for a catalytic mechanism from combinatorial antibody libraries. *Proc. Natl. Acad. Sci. USA* **91**, 2532-2536.

Janda, K. D., Lo, L.-C., Lo, C.-H., Sim, M.-M., Wang, R., Wong, C.-H. & Lerner, R. A. (1997) Chemical selection for catalysis in combinatorial antibody libraries. *Science* **275**, 945-948.

Janin, J. & Chothia, C. (1990) The structure of protein-protein recognition sites. *J. Biol. Chem.* **265**, 16027-16030.

Jestin, J-L, Kristensen, P. & Winter, G. (1999) *Angew. Chem. Int. Ed.* **38**, 1124-1127.

Johnsson, B., Lofas, S. & Lindquist, G. (1991) Immobilization of proteins to a carboxymethyl dextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Anal. Biochem.* **198**, 268-277.

Jones, S. & Thornton, J. M. (1996) Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **93**, 13-20.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) Imporved methods for the building of protein models in electron density maps and the location of errors in these models. *Acta Cryst.* **A47**, 110-119.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). In Sequences of Proteins of Immunological Interest, 5th ed., vol. 1, pp. 688–696, NIH, Bethesda, MD.

Kane, J. F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Curr. Opin. Biotechnol.* **6**, 494-500.

Kast, P. & Hilvert, D. (1997) 3D structural information as a guide to protein engineering

using genetic selection. *Curr. Opin. Struct. Biol.* **7**, 470-479.

Katz, B. A. (1997) Structural and mechanistic determinants of affinity and specificity of ligands discovered or engineered by phage display. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 27-45.

Keefe, L. J., Sondek, J., Shortle, D. & Lattman, E. E. (1993) The alpha aneurism: a structural motif revealed in an insertion mutant of staphylococcal nuclease. *Proc. Natl. Acad. Sci. USA* **90**, 3275-3279.

Korodi, I., Asboth, B. & Polgar, L. (1986) Disulfide bond formation between the active-site thiol and one of the several free thiol groups of chymopapain. *Biochemistry* **25**, 6895-6900.

Kunkel, T. A., Bebenek, K. & McClary, J. (1991) Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol* **204**, 125-139.

Kunkel, T. A., Bebenek, K. & McClary, J. (1991) *Methods Enzymol.* **204**, 125-139.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367–382.

Lasky, L. A. & Dowbenko, D. J. DNA sequence analysis of the type-common glycoprotein-D genes of herpes simplex virus types 1 and 2 (1984). *DNA* **3**, 23–29.

Li, B., Tom, J. Y., Oare, D., Yen, R., Fairbrother, W. J., Wells, J. A. & Cunningham, B. C. (1995) Minimization of a polypeptide hormone. *Science* **270**, 1657-1660.

Light, J. & Lerner, R. A. (1995) Random mutagenesis of staphylococcal nuclease and phage display selection. *Bioorg. Med. Chem.* **3**, 955-967.

Lilova, A., Kleinschmidt, T. & Nedkov, P. (1987) Methionine residue accessibility in native subtilisin DY. *Biol. Chem. Hoppe Seyler* **368**, 513-519.

Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K. & Wilson, I. A. (1996) Functional mimicry of a

protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. *Science* **273**, 464-471.

Lorsch, J. R. & Szostak, J. W. (1994) In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature* **371**, 31-36.

Lowe, G. (1976) The cysteine proteinases. *Tetrahedron* **32**, 291-302.

Lowman, H. B. (1997). Phage Display of Peptide Libraries on Protein Scaffolds. In Methods in Molecular Biology: Combinatorial Peptide Libraries (in the press).

Lowman, H. B. & Wells, J. A. (1993) Affinity maturation of human growth hormone by monovalent phage display. *J. Mol. Biol.* **234**, 564-578.

Lowman, H. B. (1998) in *Methods in Molecular Biology, vol. 87: Combinatorial Peptide Library Protocols,* ed. Cabilly, S. (Humana Press, Totowa), pp. 249-264.

Lowman, H. B., Bass, S. H., Simpson, N. & Wells, J. A. (1991). Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry,* **30**, 10832–10838.

Lowman, H. B., Bass, S. H., Simpson, N. & Wells, J. A. (1991) Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* **30**, 10832-10838.

Lowman, H. B. (1997) Bacteriophage display and discovery of peptide leads for drug development. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 401-424.

Maenaka, K., Furuta, M., Tsumoto, K., Watanabe, K., Ueda, Y. & Kumagai, I. (1996) A stable phage-display system using a phagemid vector: phage display of hen egg-white lysozyme (HEL), Escherichia coli alkaline, phosphatase, and anti-HEL monoclonal antibody, HyHEL10. *Biochem. Biophys. Res. Comm.* **218**, 682-687.

Meldal, M. & Breddam, K. (1991) Anthranilamide and nitrotyrosine as a donor-acceptor pair in internally quenched fluorescent substrates for endopeptidases: multicolumn peptide synthesis of enzyme substrates for subtilisin Carlsberg and pepsin. *Anal. Biochem.* **195**, 141-147.

Miller, S. (1990). Protein-protein recognition and the association of immunoglobulin constant domains. *J. Mol. Biol.* **216**, 965–973.

Morihara, K. & Oka, T. (1977) A kinetic invenstigation of subsites S1' and S2' in alpha-chymotrypsin and subtilisin BPN'. *Arch. Biochem. Biophys.* **178**, 188-194.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) Refinement of macromolecular structures by maximum-likelihood method. *Acta Cryst.* **D53**, 240-255.

Natatsuka, T., Sasaki, T. & Kaiser, E. T. (1987) Peptide segment coupling catalyzed by the semisynthetic enzyme thiolsubtilisin. *J. Am. Chem. Soc.* **109**, 3808-3810.

O'Shea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide "velcro": design of a heterodimeric coiled coil. *Curr. Biol.* 3, 658–667.

Otwinowski, Z. (1993). In*Data Collection and Processing*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), p. 56, SERC Daresbury Laboratory, Warrington, UK.

Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L. & Bryan, P. N. (1989) Large increases in general stability for subtilisin BPN' through incremental changesin the free energy of unfolding. *Biochemistry* **28**, 7205-7213.

Parmley, S. F. & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**, 305-318.

Patten, P. A., Howard, R. J. & Stemmer, W. P. (1997) Applications of DNA shuffling to pharmaceuticals and vaccines. *Curr. Opin. Biotech.* **8**, 724-733.

Pearce, K. H. Jr., Ultsch, M. H., Kelley, R. F., de Vos, A. M. & Wells, J. A. (1996) Structural and mutational analysis of affinity-inert contact residues at the growth hormone-receptor interface. *Biochemistry* **35**, 10300-10307.

Pearce, K. H. Jr., Cunningham, B. C., Fuh, G., Teeri, T. & Wells, J. A. (1999) Growth hormone binding affinity for its receptor surpasses the requirements for cellular activity. *Biochemistry* **38**, 81-89.

Pedersen, H., Holder, S., Sutherlin, D., Schwitter, U., King, D. S. & Schultz, P. G. (1998) A method for directed evolution and functional cloning of enzymes. *Proc. Natl. Acad. Sci. USA* **95**, 10523-10528.

Perona, J. J. & Craik, C. S. (1995) Structural basis of substrate specificity in the serine proteases. *Protein Sci.* **4**, 337-360.

Peterson, E. B. & Hilvert, D. (1995) Nonessential active site residues modulate selenosubtilisin's kinetic mechanism. *Biochemistry*, **34**, 6616-6620.

Philipp, M. & Bender M. L. (1983) Kinetics of subtilisin and thiolsubtilisin. *Mol. Cell. Biochem.* **51**, 5-32.

Picken, R. N., Mazaitis, A. J., Maas, W. K., Rey, M. & Heyneker, H. (1983). Nucleotide sequence of the gene for heat-stable enterotoxin II of *Escherichia coli. Infect. Immun.* **42**, 269-275.

Quemeneur, E., Moutiez, M., Charbonnier, J. B. & Menez, A. (1998) Engineering cyclophilin into a proline-specific endopeptidase. *Nature*, **391**, 301-304.

Rheinnecker, M., Eder, J., Pandey, P. S. & Fersht, A. R. (1994) Variants of subtilisin BPN' with altered specificity profiles. *Biochemistry* **33**, 221-225.

Richards, F. M. & Lim, W. A. (1993) An analysis of packing in the protein folding problem. *Q. Rev. Biophys.* **26**, 423-498.

Ridgway, J. B. B., Presta, L. G. & Carter, P. (1996). "Knobs-into-holes" engineering of antibody C_H3 domains for heavy chain heterodimerization. *Protein Eng.* **9**, 617–621.

Robey, E. A. & Schachman, H. K. (1985). Regeneration of active enzyme by formation of hybrids from inactive derivatives: implication for active sites shared between polypeptide chains of aspartate transcarbamoylase. *Proc. Natl. Acad. Sci. USA* **82**, 361–365.

Rodrigues, M. L., Presta, L. G., Kotts, C. E., Wirth, C., Mordenti, J., Osaka, G., Wong, W. L. T., Nuijens, A., Blackburn, B. & Carter, P. (1995). Development of a humanized disulfide-stabilized anti-$p185^{HER2}$ Fv-β-lactamase fusion protein for activation of a cephalosporin doxorubicin prodrug. *Cancer Res.* **55**, 63–70.

Rodrigues, M. L., Shalaby, M. R., Werther, W., Presta, L. & Carter, P. (1992). Engineering a humanized bispecific F(ab')$_2$ fragment for improved binding to T cells. *Int. J. Cancer Suppl.* **7**, 45–50.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

Schoneich, C., Zhao, F., Wilson, G. S. & Borchardt, T. T. (1993) Iron-thiolate induced oxidation of methionine to methionine sulfoxide in small model peptides. Intramolecular catalysis by histidine. *Biochim. Biophys. Acta* **1158**, 307-322.

Shalaby, M. R., Shepard, H. M., Presta, L., Rodrigues, M. L., Beverley, P. C. L., Feldmann, M. & Carter, P. (1992). Development of humanized bispecific antibodies reactive with cytotoxic lymphocytes and tumor cells overexpressing the *HER2* protooncogene. *J. Exp. Med.* **175**, 217-225.

Siezen, R. J. & Leunissen, J. A. M. (1997) Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci.* **6**, 501-523.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228,** 1315-1317.

Somers, W., Ultsch, M., de Vos, A. M. & Kossiakoff, A. A. (1994) The X-ray structure of a growth hormone-prolactin receptor complex. *Nature* **372**, 478-481.

Sorensen, S. B., Bech, L. M., Meldal, M. & Breddam, K. (1993) Mutational replacements of the amino acid residues forming the hydrophobic S4 binding pocket of subtilisin 309 from Bacillus lentus. *Biochemistry* **32**, 8994-8999.

Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J., Winter, G. & Fastrez, J. (1994a) Selection of beta-lactamase on filamentous bacteriophage by catalytic activity. *J. Mol. Biol.* **237**, 415-422.

Soumillion, P., Jespers, L., Bouchet, M., Marchand-Brynaert, J., Sartiaux, P. & Fastrez, J. (1994b) Phage display of enzymes and in vitro selection for catalytic activity. *Appl. Biochem. Biotechnol.* **47**, 175-190.

Spencer, S. A., Hammonds, R. G., Henzel, W. J., Rodriquez, H., Waters, M. J. & Wood, W. I. (1988) Rabbit liver growth hormone receptor and serum binding protein. Purification, characterization, and sequence. *J. Biol. Chem.* **263**, 7862-7867.

Srinivasan, N., Sowdhamini, R., Ramakrishnan, C. & Balaram, P. (1990) Conformations of disulfide bridges in proteins. *Int. J. Pept. Protein Res.* **36**, 147-155.

Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391.

Sundstrom, M., Lundqvist, T., Rodin, J., Biebel, L. B., Milligan, D. & Norstedt, G. (1996) Crystal structure of an antagonist mutant of human growth hormone, G120R, in complex with its receptor at 2.9 A resolution. *J. Biol. Chem.* **271**, 32197-32203.

Takagi, H., Morinaga, Y., Ikemura, H. & Inouye, M. (1988) Mutant subtilisin E with enhanced protease activity obtained by site-directed mutagenesis. *J. Biol. Chem.* **263**, 19592-19596.

Tawfik, D. S., Green, B. S., Chap, R., Sela, M. & Eshhar, Z. (1993) catELISA: A facile general route to catalytic antibodies. *Proc. Natl. Acad. Sci. USA* **90**, 373-377.

Vetter, I. R., Baase, W. A., Heinz, D. W., Xiong, J. P., Snow, S. & Matthews, B. W. (1996) Protein structural plasticity exemplified by insertion and deletion mutants in T4 lysozyme. *Protein Sci.* **5**, 2399-2415.

Vieira, J. & Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods Enzymol.* **153**, 3–11.

Walter, M. R., Windsor, W. T., Nagabhushan, T. L., Lundell, D. J., Lunn, C. A., Zauodny, P. J. & Narula, S. K. (1995) Crystal structure of a complex between interferon-gamma and its soluble high-affinity receptor. *Nature* **376**, 230-235.

Ward, W. H. J., Jones, D. H. & Fersht, A. R. (1987). Effects of engineering complementarity charged residues into the hydrophobic subunit interface of tyrosyl-tRNA synthetase. *Biochemistry* **26**, 4131–4138.

Wells, J. A. & de Vos, A. M. (1996) Hematopoietic receptor complexes. *Annu. Rev. Biochem.* **65**, 609-634.

Wells, J. A. & Estell, D. A. (1988) Subtilisin--an enzyme designed to be engineered. *Trends Biochem. Sci.* **13**, 291-297.

Wente, S. & Schachman, H. K. Shared active sites in oligomeric enzymes: model studies with defective mutants of aspartate transcarbamoylase produced by site-directed mutagenesis. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 31–35.

Widersten, M. & Mannervik, B. (1995) Glutathione transferases with novel active sites isolated by phage display from a library of random mutants. *J. Mol. Biol.* **250**, 115-122.

Wilson, D. R. & Finlay, B. B. (1998) Phage display: applications, innovations, and issues in phage and host biology. *Can. J. Microbiol.* **44**, 313-29.

Wu, P. & Brand, L. (1994) Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1-13.

Yaron, A., Carmel, A. & Katchalski-Katzir, E. (1979) Intramolecularly quenched fluorogenic substrates for hydrolytic enzymes. *Anal. Biochem.* **95**, 228-235.

Zhang, B. & Cech, T. R. (1998) Peptidyl-transferase ribozymes: trans reactions, structural characterization and ribosomal RNA-like features. *Chem. Biol.* **5**, 539-553.

Zhao, H. & Arnold, F. H. (1997) Combinatorial protein design: strategies for screening protein libraries. *Curr. Opin. Struct. Biol.* **7**, 480-485.

Zhu, Z., Presta, L. G., Zapata, G. & Carter, P. (1997). Remodeling domain interfaces to enhance heterodimer formation. *Protein Sci.* **6**, 781-788.