High Affinity Fn3 Engineering:
Recursive Mutagenesis, Loop Length Diversity, and Loop Shuffling

Benjamin J. Hackel[1], Atul Kapila[2], and K. Dane Wittrup[1,3]

Chemical Engineering[1], Biology[2], and Biological Engineering[3]

Massachusetts Institute of Technology, Cambridge, MA

The tenth type III domain of human fibronectin (Fn3) is a small (10 kDa), stable protein with a beta-sandwich fold and three solvent-exposed loops. It has been demonstrated as an effective protein scaffold to engineer high affinity binders to multiple targets using mRNA display, phage display, and yeast two-hybrid methods. Yet, while variability in the length of complementarity-determining regions is recognized as a critical factor in antibody diversity, past Fn3 libraries have had fixed loop lengths. The current work addresses if loop length diversity can improve the affinity and recognition capacity of the Fn3 scaffold. A second, broader question is if recursive mutagenesis of selected populations, already a necessary component of *in vitro* display methods through PCR, enables rapid isolation of high affinity binders using yeast surface display libraries.

Oligonucleotides of multiple lengths containing NNB degenerate codons were used to diversify the length and amino acid composition of the loops. The yeast surface display library had 2x10^7 Fn3 clones with four possible lengths of each loop. Clones that bound lysozyme were selected by multiple rounds of fluorescence-activated cell sorting. The enriched population was diversified every two to three rounds by error-prone PCR, loop shuffling, and homologous recombination using a simple one-day protocol. Isolated clones from each stage of the selection were sequenced and characterized in terms of affinity and association and dissociation kinetics.

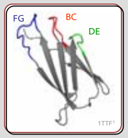
Multiple diverse clones with picomolar equilibrium dissociation constants were identified, with 2.8 ± 0.5 pM as the highest affinity. The high affinity clones are monomeric and stable (T_m = 52-58°C). The majority of high affinity clones had BC and DE loops one amino acid shorter than wild-type suggesting that loop length diversity is a valuable element of Fn3 library design. Isolation of such high affinity clones from a relatively small initial library was enabled by affinity maturation via continued selection and diversification as demonstrated by parallel selections without mutagenesis and shuffling. The recurrent population diversification should be advantageous to protein engineering in general.

Background

Tenth type III domain of human fibronectin (Fn3)

Biophysical properties

- β-sandwich with 3 solvent-exposed loops (BC, DE, FG)
- small: 94 amino acids, 10 kDa
- stable: T_m = 84°C
- Cys-free: 1-50 mg/L yield in *E. coli*, intracellular use
- soluble: 15 mg/mL



Functional binding protein

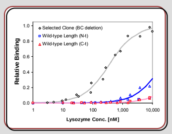
- native Fn3: critical in integrin binding
- Koide: diversify two loops → K_d~nM-μM [2-5]
- Phylos: diversify three loops → K_d ~ pM-nM [6,7]
- however previous high affinity binders are significantly destabilized [7]

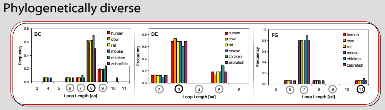
Effective in numerous applications

- inhibitors in cell culture; extra- and intracellular [4,5]
- immobilized in protein arrays [6]
- labeling reagents (Western blot, flow cytometry) [3,5]
- clinical trial of therapeutic (Adnexus)

Loop Length Diversity

- CDR length is a key element of antibody diversity [8,9]
  - Affinity maturation is improved with inclusion of loop length diversity [10,11]
  - CDR length correlates with ligand type [12,13]
  - Amino acid insertions/deletions are stably tolerated [14]
- Engineered binders exhibit undesigned non-wt lengths
  - TNFα: 2 non-wt BC and 2 non-wt FG [6]
  - Lysozyme: multiple with non-wt BC [15]
    - reversion to wt significantly decreases affinity
  - CEA: non-wt BC
- Fn3 loops tolerate Gly_4 insertions [16]
- Phylogenetically diverse





- yet past Fn3 libraries have maintained fixed loop lengths

→ **include loop length diversity in library design**

Recursive Mutagenesis

- Yeast surface display
  - \+ flow cytometry + stability engineering
  - – yeast transformation limits library size (10^7-10^9)
- Alternative display formats also substantially undersample sequence space ($20^{23} = 10^{30}$)
  - phage: 10^5-10^{10}, mRNA: 10^{11}-10^{13}, ribosome: 10^{11}-10^{13}
- *in vitro* formats incorporate recursive mutagenesis via PCR
  - affinity maturation during lead isolation

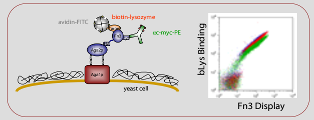
→ **include recursive mutagenesis in yeast surface display**

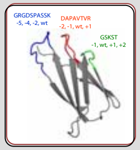
- extensive search of sequence space
- explore sequence space in the vicinity of 1000s of lead molecules

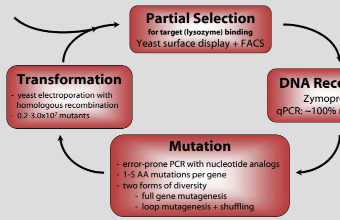
Protein Engineering

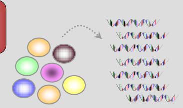
Fn3 library design

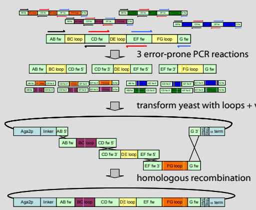
- Construct w/ overlap extension PCR
- NNB diversity
- 3 loops diversified, 4 lengths each
- 2.3x10^7 full-length clones







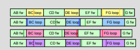




genes with mutated + shuffled loops

Gene Diversity

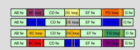
1-5 AA mutations in loops and framework



- Maintain cooperative loops
- Identify potentially beneficial framework mutations
- Make smaller changes in sequence

Loop Diversity

1-5 AA mutations in loops + shuffled loops

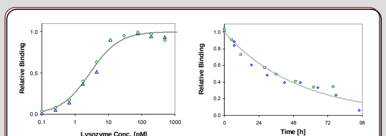


- Eliminate neutral/detrimental loops
- Focus mutation on critical loop amino acids
- Make larger changes in sequence

Results

Three clones with 3 pM affinity

Clone	K_d [pM]	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]	$\tau_{1/2}$ [h]
7.5.1	3.1 ± 0.6	1.3x10^6	2.5 ± 0.2x10^{-6}	78 ± 8
8.5.2	2.9 ± 1.4	n/d	3.2 ± 0.1x10^{-6}	60 ± 1
8.5.7	2.8 ± 0.5	n/d	5.4 ± 0.6x10^{-6}	36 ± 4

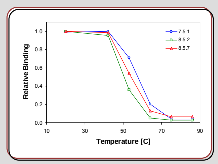


Highest affinity clone (8.5.7)

Monomeric and stable

Clone	Monomer	T_m [°C]
7.5.1	99%	58
8.5.2	80%*	52
8.5.7	93%	54

* potential disulfide bond



Results (Cont'd)

Impact of design and engineering elements

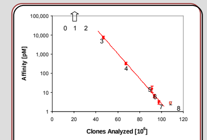
Loop length diversity

Non wild-type loop lengths occur in multiple loops of highest affinity binders

Clone	BC Loop	DE Loop	FG Loop	K_d [pM]
	Loop Sequences			
7.5.1	RGYPWAT	GVTN	RVGRTFDTPG	3.1 ± 0.6
8.5.2	RGCPWAI	GVTN	RVGRMLCAPG	2.9 ± 1.4
8.5.7	RDRPWAI	GVTN	RLSIVPYA	2.8 ± 0.5
length (to wt)	-1	-1	-2, wt	

Recursive mutagenesis

Efficient maturation:
3 pM from 1x10^8 clones
1.5x10^7 clones → 10x improvement



round-by-round affinity maturation

Loop shuffling

BC motif in initial library with ineffective DE, FG
DE, FG motifs originally matched with other BC

Point mutations

Important mutations in loops and framework

Round	BC Loop	DE Loop	FG Loop	'Framework'	K_d [pM]
	Amino Acid Sequences				
0	RDCPWAT	WTPVCF	SSQRGCM	none	>>100,000
1	SLDNQAN	GQSD	RCEPSRNSAV	none	>100,000
2	same as round 1				
3	SLDNQAN	GVTN	RVGRMLDTPG	P44S, V50M	7600 ± 1100
4	SLDNQAK	GATN	RCKPFRNSAV	P44S, V50M, T97I	330 ± 50
5	RDCPWAL	GVTN	RVGRMSCTSG	V1A, S1P, T14A, R33G, P44L, V50M	16 ± 6
6	RGCPWAI	GVTN	RVGRMLCTPG	P15S, R33G, N42S, P44L, V50M, K63E	6.6 ± 1.3
7	RGYPWAT	GVTN	RVGRTFDTPG	P15S, R33G, T35I, P44L, V50M	3.1 ± 0.6
8 (A)	RGCPWAI	GVTN	RVGRMLCAPG	R33G, I34V, N42S, P44L, V45A, V50M, K63E	2.9 ± 1.4
8 (B)	RDRPWAI	GVTN	RLSIVPYA	D3G, L18I, R33G, N42S, P44L, V50M, Y73H, N91T	2.8 ± 0.5

highest affinity clone from each sub-library

Conclusions

- Shape (loop length) is an important diversity element
- Modified affinity maturation method enables engineering of low pM binders from a small library
  - Inclusion of recursive mutagenesis in yeast surface display
  - Combination of gene mutagenesis and loop mutagenesis/shuffling

Future Work

- Engineer high affinity Fn3 binders to other targets
- Investigate Fn3 : lysozyme interaction
  - impact of each loop
  - impact of framework mutations
  - Is 3 pM the affinity limit?
- Stability engineering of high affinity binders

Acknowledgements

This work was supported by CA96504, CA101830, and a National Defense Science and Engineering Graduate fellowship (BJH). We thank Dasa Lipovsek for many useful discussions. We appreciate assistance from the MIT Flow Cytometry Core facility and the MIT Biophysical Instrumentation facility.

1. Main, *et al.* *Cell* (1992)
2. Koide, *et al.*, *J. Mol. Biol.* (1998)
3. Karatan, *et al.*, *Chem. Biol.* (2004
4. Koide, *et al.*, *PNAS* (2002
5. Richards, *et al.*, *J. Mol. Biol.* (2003)
6. Xu, *et al.*, *Chem. Biol.* (2002)
7. Parker, *et al.*, *Prot. Engr.* (2005)
8. Rock, *et al.*, *J. Exp. Med.* (1994)
9. Ohlin and Borrebaeck, *Mol. Imm.* (1996)
10. Lamminmaki, *et al.*, *J. Mol. Biol.* (1999)
11. Lee, *et al.*, *J. Mol. Biol.* (2004)
12. Collis, *et al.*, *J. Mol. Biol.* (2003)
13. Persson, *et al.*, *J. Mol. Biol.* (2006)
14. Lantto and Ohlin, *J. Biol. Chem.* (2002)
15. Lipovsek, *et al.*, *J. Mol. Biol.* submitted
16. Batori, *et al.*, *Prot. Engr.* (2002)