ACS SyntheticBiology

Letter

pubs.acs.org/synthbio

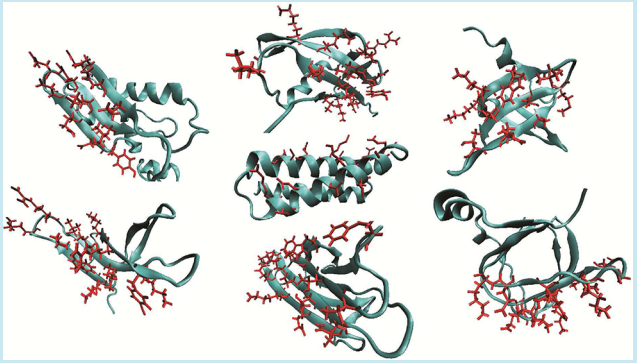
Scaffold Diversification Enhances Effectiveness of a Superlibrary of Hyperthermophilic Proteins

Mahmud Hussain,† Nimish Gera,† Andrew B. Hill, and Balaji M. Rao*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina

Supporting Information

ABSTRACT: The use of binding proteins from non-immunoglobulin scaffolds has become increasingly common in biotechnology and medicine. Typically, binders are isolated from a combinatorial library generated by mutating a single scaffold protein. In contrast, here we generated a "superlibrary" or "library-of-libraries" of 4×10^8 protein variants by mutagenesis of seven different hyperthermophilic proteins; six of the seven proteins have not been used as scaffolds prior to this study. Binding proteins for five different model targets were successfully isolated from this library. Binders obtained were derived from five out of the seven scaffolds. Strikingly, binders from this modestly sized superlibrary have affinities comparable or higher than those obtained from a library with 1000-fold higher sequence diversity but derived from a single stable scaffold. Thus scaffold diversification, i.e., randomization of multiple different scaffolds, is a powerful alternate strategy for combinatorial library construction.



KEYWORDS: *hyperthermophilic protein scaffolds, yeast surface display, protein engineering, alternate scaffold, mRNA display*

Binding proteins derived from non-immunoglobulin scaffolds offer several advantages over antibodies such as smaller size, higher stability, and ease of recombinant expression at high yields in bacterial systems. Indeed, the generation of binding proteins from several different protein scaffolds has been reported.[1,2]

Binders are typically isolated from a combinatorial library generated by randomization of a surface-accessible region on the scaffold. The likelihood of obtaining binders to any given target is greatest when the scaffold can yield mutants that collectively have large topological diversity; this is the case with immunoglobulins when the amino acid composition of complementarity determining regions (CDRs) is varied. We hypothesized that simultaneous randomization of multiple proteins is a powerful strategy for constructing a topologically diverse non-immunoglobulin scaffold library that is capable of generating binders to a wide spectrum of targets. Accordingly, we generated a "superlibrary" by randomizing secondary structure elements of seven different hyperthermophilic proteins; six of these have not been previously used as scaffolds for generating binding proteins. We reasoned that highly stable proteins are more likely to be tolerant to mutation and retain their native structure. Therefore, hyperthermophilic proteins are excellent candidates for scaffolds; libraries derived from randomizing these will contain fewer misfolded mutants.[3,4]

We screened the superlibrary for binders to five different model targets; our results confirm that binders to diverse target species can indeed be isolated. We further assessed the benefits of scaffold diversification by simultaneously isolating binders from the superlibrary as well as a library of Sso7d mutants with 1000-fold higher sequence diversity. The Sso7d protein from *Sulfolobus solfataricus* has been previously shown to be a versatile scaffold that can yield binders to a wide range of target species.[4] Notably, affinities of binders obtained using these approaches are comparable and underscore the effectiveness of using multiple scaffolds.

RESULTS AND DISCUSSION

Construction of a Superlibrary of Alternate Scaffolds Using Yeast Surface Display. We constructed a combinatorial superlibrary by randomizing specific regions on seven different proteins: Tm0487 and Tm1112 from *Thermotoga maritima*; Sso7d, Sso6901, and the microtubule interacting and transport (MIT) domain of Sso0909 from *Sulfolobus solfataricus*; the chitin binding domain (ChBD) of Chitinase A from *Pyrococcus furiosus*; and the N-terminal domain of the Ph1500 protein from *Pyrococcus horikoshii*. Candidate proteins from hyperthermophilic organisms were chosen on the basis of two criteria, a known NMR or crystal structure that allows careful selection of surface-accessible residues for mutagenesis and overall length of ~100 amino acids or less. The seemingly arbitrary restriction on the overall length of scaffold proteins was necessary to ensure that all individual scaffold libraries could be generated using one or two oligonucleotides containing degenerate codons. Six of the seven proteins chosen,

Received: April 11, 2012
Published: June 15, 2012

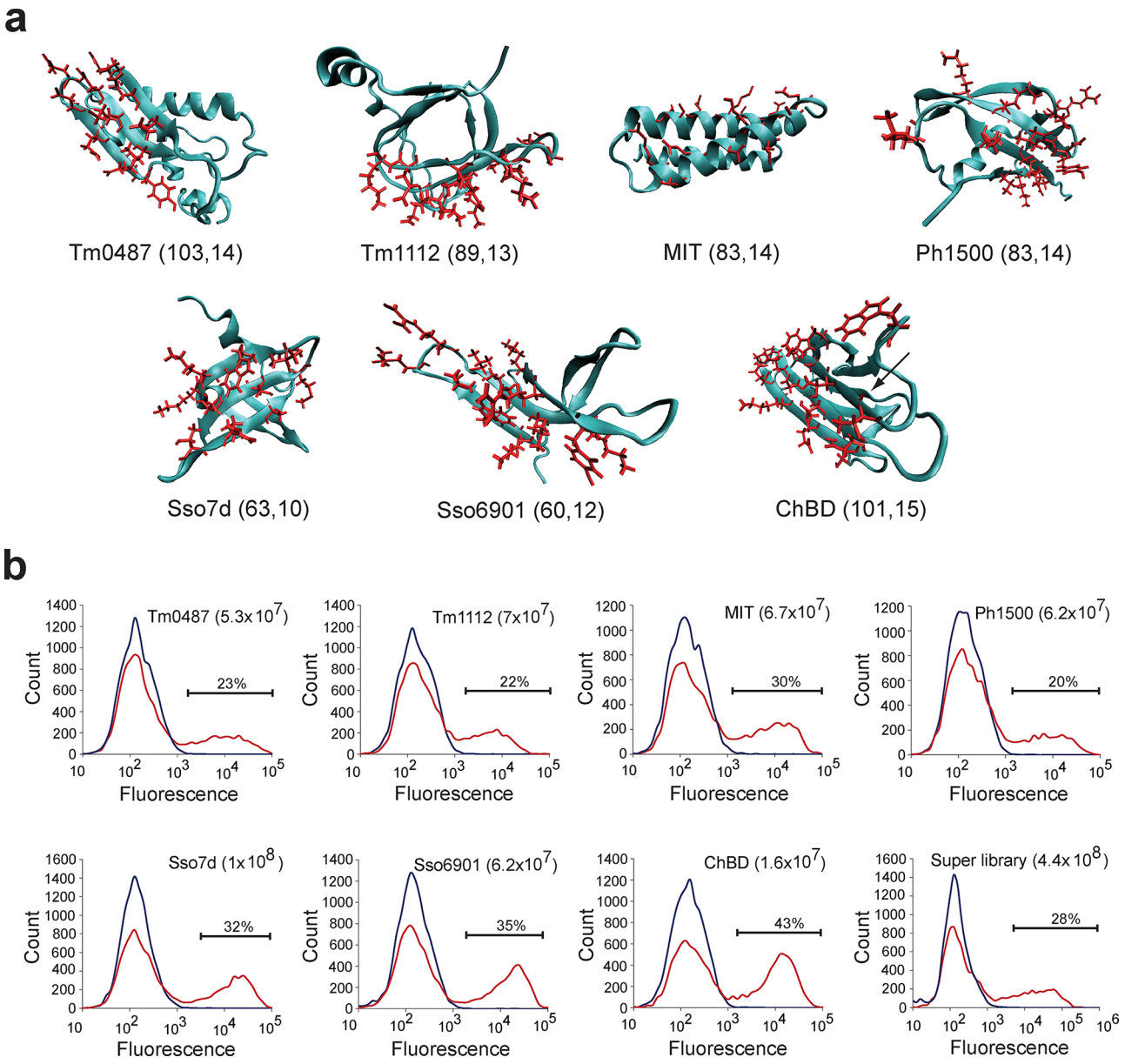


Figure 1. Construction of a superlibrary of hyperthermophilic scaffolds. (a) Scaffolds used to construct the superlibrary. Residues chosen for randomization are shown in licorice representation; Visual Molecular Dynamics (VMD) v1.8.7 was used to generate this figure. Also see Supplementary Table S1. For each scaffold, the total number of amino acid residues in the wild-type protein, followed by the number of residues randomized, is shown in parentheses. An arrow indicates the site of an inadvertent insertion of a threonine residue during construction of the ChBD library. (b) Analysis of cell surface protein expression for individual scaffold libraries and the combined superlibrary; numbers in parentheses indicate library diversity. Yeast cells displaying mutant proteins as cell surface fusions were labeled with an anti-c-myc antibody followed by a secondary antibody conjugated with Alexa Fluor 488 (red histograms) or the fluorescent secondary antibody alone (purple histograms). Cells were subsequently analyzed using flow cytometry. Numbers above the horizontal bar show the approximate percentage of yeast cells in each library expressing the c-myc epitope tag and therefore, by inference, full-length proteins as cell surface fusions.

all proteins except Sso7d, have not been used as scaffolds for generating binding proteins prior to this study; we have previously isolated binding proteins for a wide spectrum of target species from a library of 10^8 Sso7d mutants.[4]

Combinatorial libraries of scaffold proteins were generated by randomizing 10−15 residues on each scaffold, as shown in Figure 1a and Supplementary Table S1. In general, surface-accessible residues were chosen for randomization; where available, additional data in the literature were used to guide the choice of residues (see Supplementary Methods). Yeast display libraries of 10^7−10^8 mutants for each scaffold were generated (Figure 1a). Mutant proteins are expressed as a C-terminal fusion to the yeast cell wall protein Aga2, and are flanked by an N-terminal HA and C-terminal c-myc epitope tags, respectively.[5] Subsequently, individual libraries were combined to obtain a superlibrary with an overall diversity of ~4.4×10^8 mutants. This yeast superlibrary can be propagated indefinitely in culture or frozen. Indirect immunofluorescent labeling of the c-myc epitope tag shows that 20−43% of yeast cells in the individual libraries express full-length fusions (Figure 1b). Overall, 28% of yeast cells in the superlibrary express full-length scaffold variants.

The ChBD library was constructed using oligonucleotides with degenerate trinucleotide phosphoramidites coding for 20

dx.doi.org/10.1021/sb300029m | ACS Synth. Biol. 2013, 2, 6−13

Table 1. Protein Sequences of Mutants Obtained from the Superlibrary or the Sso7d mRNA display library[a]

Scaffold/mutant	Sequence
	.30.....40.....50.....60.....70..
MIT-WT	EDAITYYKKAIEVLSQIIVLYPESVARTAYEQMINEYKKRISYLE
MIT-rIgG	GDAISYYWKAIVVLQRIIVLYPESVATGAYLQMIGEYARRIDYLN
	......30.....40.....50..
Sso6901-WT	KVWALAPKGRKGVKIGLFKDPETGKYFRHK
Sso6901-BCP	LVQARAPFGRKGVKRGLFRDPETGKAFFHL
Sso6901-BNP32	SVYAFAPCGLKGSKWGSFLDPETGKYFDHV
	20.....30.....40...
Sso7d-WT	KKVWRVGKMISFTYDLGGGKTGRGA
Sso7d-HEL	CFVFRWGKCICFDYDLGGGKQGSGC
Sso7d- BCP	NPVVRYGKLIFFAYDLGGGKLGAGW
Sso7d-Fluorescein-2	KFVLRPGKAILFYYDLGGGKYGFGL
Sso7d-Fluorescein-3	LKVFRIGKVIFFRYDLGGGKFGYGY
Sso7d-BNP32-2	INVNRGGKFIRFTYDLGGGKFGSGR
From mRNA display library	
Sso7d-rIgGm	YRVFRSGKTIFFRYDLGGGKLGVGI
Sso7d-rIgG-2	YWVRRHGKSITFQYDLGGGKNGLGF
Sso7d-rIgG-3	QLVRRRGKRITFRYDLGGGKKGVGY
Sso7d-BNP32m	YCVKRSGKKIRFFYDLGGGKRGIGT
Sso7d-BNP32-3	ARVWRVGKRILFGYDLGGGKLGIGR
Sso7d-BNP32-4	LYVYRIGKRIIFAYDLGGGKVGVGW
Sso7d-BNP32-5	NVVSRIGKTITFLYDLGGGKMGSGV
Sso7d-BNP32-6	SWVFRRGKYIIFAYDLGGGKGGHGK
Sso7d-BNP32-7	QAVSRIGKHINFKYDLGGGKQGSGC
	...50.....60.....70.....80
Tm1112-WT	KVEVTTEDGKKYVIEKGDLVTFPKGLRCRWKVLE
Tm1112-BNP32[b]	FVCVLTMDGKTYGID----------
	.267.....277.....287...I....297.....307..
ChBD-WT[c]	SLEVKVNDWGSGAEYDVTLNLDGQYDWTTVKVKLAPGATVGSFWSAN
	.313.....323.....333..
	KQEGNGYVIFTPVSWNKGPTATFGFI
	.267.....277.....287...I....297.....307..
ChBD-rIgG[c]	PLRVRVVDLGSGAWYLVLLHLDGQYDWTTVKVKLAPGATVGSFYSAN
	.313.....323.....333..
	KQEGNGYVIFTPVSQNKGPQAWFEFE

[a]Corresponding wild-type (WT) sequences are shown as a reference. Mutated residues are in bold font and underlined. [b]A truncated Tm1112 mutant was obtained due to the introduction of a stop codon at K62. [c]A threonine insertion was inadvertently introduced between W292 and T293 (indicated by I) during oligonucleotide synthesis; this insertion is shown in bold font and italicized.

dx.doi.org/10.1021/sb300029m | *ACS Synth. Biol.* 2013, 2, 6−13

amino acids. Note that all stop codons are eliminated in this approach.[6] Yet, only 43% of cells in the ChBD yeast library expressed full-length cell surface fusions. In contrast, 30% of the Sso7d library expresses full-length fusions even though only ~30% of the library consists of full-length Sso7d variants due to frameshift mutations and stop codons introduced by degenerate NNN codons used in library construction.[4] It has been previously proposed that the quality control mechanism in the yeast endoplasmic reticulum allows cell surface expression of only correctly folded proteins.[7] Therefore, assuming a similar frequency of frameshift mutations in the ChBD and Sso7d libraries, it is likely that a larger fraction of mutants in the Sso7d library are correctly folded relative to the ChBD library, although there is some evidence of yeast surface expression of misfolded proteins or molten globules.[8] Indeed, an amino acid insertion was inadvertently introduced in a β-sheet in the ChBD scaffold, resulting in likely destabilization of all ChBD mutants (Figure 1a). Moreover, a ChBD mutant isolated from the superlibrary has the lowest thermal stability among all proteins analyzed in this study (*vide infra*). Taken together, these data are consistent with the idea that highly stable scaffold proteins yield combinatorial libraries containing a larger fraction of correctly folded proteins.

Isolation of Binders To Model Targets from the Superlibrary. Using magnetic selection followed by fluorescence activated cell sorting (FACS) as previously described,[5] we successfully isolated binders for five different model targets from the superlibrary. The targets chosen included a small organic molecule (fluorescein), a 12 amino acid peptide from the C-terminus of the β-catenin protein (BCP), the 32 amino acid Brain Natriuretic Peptide32 (BNP32) that contains a disulfide bridge, hen egg lysozyme (HEL), and immunoglobulin G from rabbit (rIgG). These targets represent species with a range of molecular weights; also specificity of rIgG binders could be rigorously assessed by evaluating their binding to other closely related immunoglobulins (*vide infra*).

The pool of mutants with the highest affinity for the targets, obtained after multiple rounds of FACS, included proteins derived from five of the seven scaffolds used to construct the superlibrary: Sso7d, Sso6901, MIT, Tm1112, and ChBD (Table 1). Binding proteins from multiple scaffolds were obtained in the pool of highest affinity binders for all targets, with the exception of HEL. Interestingly, single distinct Sso7d mutant in the pool of highest affinity binders from the superlibrary is identical to the mutant that was previously isolated by screening the Sso7d library.[4] Also, the pool of binders for BNP32 from sorts contained a truncated Tm1112 mutant.

We measured the equilibrium dissociation constants (K_D) of the binding interaction between mutant proteins and their targets for a selected subset of mutant proteins, using yeast cell surface titrations.[5] Note that multiple previous reports have shown the consistency between K_D values estimated using yeast surface titrations and those obtained using soluble proteins.[9,10] Since rIgG is a dimeric molecule, and our estimates for K_D are based on the assumption of a monovalent binding isotherm. K_D estimates from yeast surface titrations are likely influenced by the avidity effect in the case of rIgG; a single rIgG molecule may bind to two cell surface fusions. Therefore, K_D values for rIgG binders were determined by ELISA. The estimated binding affinities and their associated 68% confidence intervals, analogous to the commonly reported standard deviations from triplicate measurements, are shown in Table 2.

Table 2. K_D Estimates for Selected Mutants[a]

Target molecule	Mutant-scaffold	Estimated K_D (nM) (68% confidence interval)
Fluorescein[b]	Sso7d-Fluorescein-2	879 (515−1700)
BCP[b]	Sso6901-BCP	3297 (2120−5300)
BNP32[b]	Tm1112-BNP32	653 (565−755)
	Sso6901-BNP32	2100 (1560−2850)
	Sso7d-BNP32m[d]	2696 (2000−3650)
HEL[b]	Sso7d-Lysozyme	349[e] (225−540)
rIgG[c]	MIT-rIgG	1156 (830−2050)
	Sso7d-rIgGm[d]	2638 (1855−4170)

[a]In each case, K_D values were estimated using data from three different experiments. See Figure 2 and Supplementary Figure S1 for titration curves. [b]Yeast surface titrations were used to estimate K_D. [c]ELISA was used to estimate K_D. [d]These binders were isolated from the mRNA display library of Sso7d mutants. [e]The HEL binder isolated from the superlibrary was identical to a mutant previously isolated.[4]

Our analysis shows that the binding affinities of mutants obtained from the superlibrary are in the few hundred nanomolar to micromolar range. Note that an affinity maturation step has not been performed on these mutants. Further increases in affinity may be achieved by additional rounds of mutagenesis and screening. However, it is important to note that several applications such as the design of affinity ligands for chromatographic separations do not require binding proteins with high affinity. For instance, the K_D for the binding interaction between protein G and IgG is ~100 nM.[11]

Role of Scaffold versus Sequence Diversity in Combinatorial Libraries. A significant fraction of the superlibrary is composed of a single scaffold; Sso7d mutants make up 10^8 out of the overall 4.4×10^8 mutants in the superlibrary. Nevertheless, the pool of binders with the highest affinity for each target was found to be derived from a distinct subset of scaffolds. For example, the "best" binders for rIgG were derived from MIT and ChBD, whereas binders to BNP32 were derived from Tm1112, Sso7d, and Sso6901 (Table 1). These results suggested that specific scaffolds may be more suited than others to generate binders for a given target.

However, the overall diversity of a yeast surface display library is restricted by the transformation efficiency in yeast; the highest reported diversities are on the order of $\sim 10^{9-10}$ mutants, whereas the theoretical diversity generated by randomizing 10 residues is $\sim 10^{13}$. The limited diversity of our individual scaffold libraries raises the following question: is the apparent propensity of a particular scaffold or subset of scaffolds to generate binders for a given target simply an artifact of library size limitation? In other words, can a library with very high sequence diversity but generated from a single versatile scaffold such as Sso7d yield binders with affinities higher than those obtained from the superlibrary with lower sequence diversity but derived from multiple scaffolds?

To address this question, we screened a library of $\sim 5 \times 10^{11}$ Sso7d mutants for binders to two targets: rIgG and BNP32. A combination of mRNA display[12] and yeast display was used. After multiple rounds of FACS, six and three distinct clones were obtained in the case of binders to BNP32 and rIgG, respectively (Table 1). We estimated the affinities of rIgG- and BNP32-binding mutants using ELISA and yeast cell surface titrations. Strikingly, binding affinities of mutants isolated from a 5×10^{11} library of Sso7d mutants are lower, or at best comparable, to those obtained from the superlibrary with overall diversity lower by 3 orders of magnitude ($\sim 4 \times 10^8$)

dx.doi.org/10.1021/sb300029m | *ACS Synth. Biol.* 2013, 2, 6−13

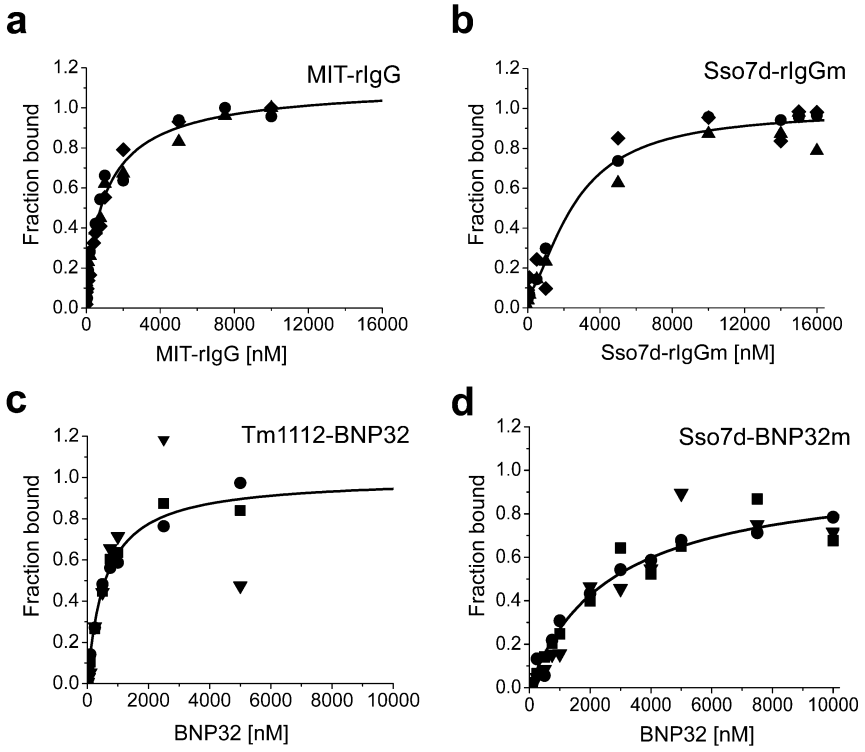


Figure 2. Comparison of binders from superlibrary and Sso7d mRNA display library. Binding curves used to estimate K_D for binding interaction of rIgG with (a) MIT-rIgG and (b) Sso7d-rIgGm, obtained by ELISA. Ninety-six-well plates were coated with rIgG followed by labeling with varying concentrations of 6xHis-tagged recombinant MIT-rIgG or Sso7d-rIgGm proteins. The fraction of immobilized rIgG bound by soluble mutant proteins is plotted as a function of mutant protein concentration. (c, d) Yeast cell surface titration plots used to estimate the apparent K_D for binding interaction of BNP32 with Tm1112-BNP32 and Sso7d-BNP32m, respectively. Yeast cells displaying mutant proteins as cell surface fusions were labeled with biotinylated BNP32, followed by streptavidin-phycoerythrin (SA-PE). Subsequently, cells were analyzed using flow cytometry. The fraction of cell surface displayed mutant proteins bound by BNP32 is plotted as a function of BNP32 concentration. In all plots, different symbols denote data from independent experiments.

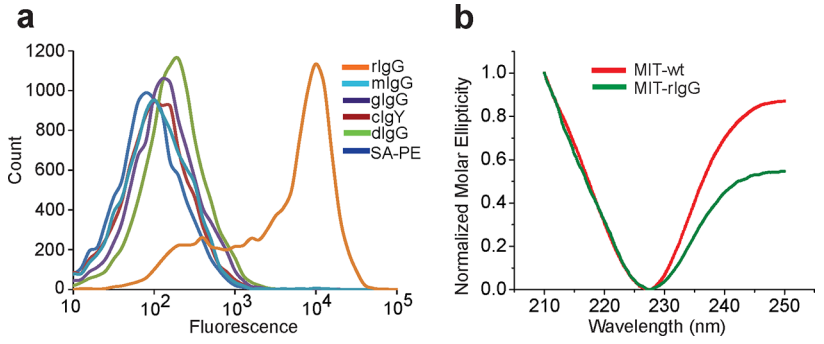


Figure 3. Characterization of MIT-rIgG. (a) Analysis of binding specificity of MIT-rIgG. Yeast cells displaying MIT-rIgG as cell surface fusions were labeled with 1 μM rIgG or an equivalent concentration of the non-target immunoglobulin species donkey IgG (dIgG), goat IgG (gIgG), chicken IgG (cIgY), and mouse IgG (mIgG). Fluorescence histograms confirm the high specificity of MIT-rIgG binding to rIgG; binding to other non-target immunoglobulins is insignificant. (b) CD spectra of wild-type MIT (MIT-WT) and MIT-rIgG mutants in 20 mM sodium phosphate buffer, pH 7.4. The baseline corrected molar ellipticity (θ) was normalized as $\bar{\theta} = (\theta - \theta_{min})/(\theta_{max} - \theta_{min})$, where θ_{min} and θ_{max} are the minimum and maximum values of baseline corrected molar ellipticity.

(Figure 2, Table 2). Further, in yeast cell surface titrations for two other BNP32-binding mutants (Sso7d-BNP32-3 and Sso7d-BNP32-4 in Table 1) isolated from the 5×10^{11} library of Sso7d mutants, saturation of the binding isotherm was not attained even at 10 μM BNP32 concentration (data not shown); this indicates that these mutants have lower affinities (i.e., higher K_D) than that reported for Sso7d-BNP32m in Table 2. Thus, our results strongly suggest that scaffold diversification is advantageous for design of combinatorial scaffold libraries. A particular scaffold may be better suited than others for generating binders to a given target.

In generating our superlibrary, we chose to randomize 10−15 amino acid residues on β-strands or α-helices of various scaffold proteins; the basic flat surface topology of the randomized region on the scaffold was maintained (Figure 1a). Arguably, inclusion of additional scaffolds where loop regions are mutagenized might enhance the binding capabilities of the superlibrary. Nevertheless, it is interesting to note that the possibly subtle variations in topology and/or amino acid

dx.doi.org/10.1021/sb300029m | ACS Synth. Biol. 2013, 2, 6−13

composition of the scaffold framework may favor a particular scaffold in the context of generating binding proteins to a given target. In some ways, this is analogous to antibodies for two different targets arising from two different germ lines. Parallels can also be drawn to the introduction of loop length diversity in CDRs of antibodies[13−16] or the 10th domain of fibronectin to generate high affinity binders.[17,18]

Analysis of Specificity and Thermal Stability. Mutant proteins isolated from the superlibrary did not exhibit any significant binding to secondary reagents (data not shown). We further analyzed the specificity of MIT-rIgG by assessing its binding to immunoglobulins from species other than rabbit. Briefly, flow cytometry was used to detect the binding of yeast cells displaying MIT-rIgG to its cognate target, rIgG, or immunoglobulins from mouse, goat, chicken, and donkey. As shown in Figure 3a, MIT-rIgG shows negligible binding to non-target immunoglobulins. We attribute the high specificity of MIT-rIgG to the stringent negative selection steps employed during magnetic selection. Additionally, circular dichroism (CD) spectroscopy confirmed that the CD spectra for wild-type and MIT-rIgG are reasonably similar over the range of wavelengths from 210 to 240 nm. CD spectra of several Sso7d mutants have been shown to be similar to that of the wild-type protein.[4]

We assessed the thermal stability of proteins derived from different scaffolds in the superlibrary, using differential scanning calorimetry (DSC) or thermal denaturation studies on yeast cell surface displayed proteins. DSC was used to measure the melting temperatures (T_m) for the wild-type proteins Tm1112, Sso6901, and MIT, as well as selected mutants derived from Sso7d, Sso6901, and MIT; these proteins were recombinantly expressed in *E. coli*. For ChBD-rIgG and Tm1112-BNP32, we used yeast cell surface displayed protein to determine the temperature of half-maximal irreversible thermal denaturation ($T_{1/2}$), as previously described.[17,19] Pertinently, DSC experiments on Tm1112-BNP32 were not feasible due to the low protein yield and aggregation-prone nature of the protein under the buffer conditions tested. Also, we chose to use yeast-displayed protein to assess the thermal stability of ChBD-rIgG due to an amino acid insertion in a β-strand; we anticipated, as indeed confirmed by experiment, that this may lead to loss of protein stability.

Table 3 shows the T_m and $T_{1/2}$ values for the wild-type and mutant proteins. Corresponding thermal denaturation curves are shown in Supplementary Figure S2. Mutant proteins derived from the Sso7d, Sso6901, and MIT scaffolds show high melting temperatures (T_m = 74−93 °C). It is noteworthy here that even though these mutants have T_m's lower than those of their parent scaffolds, they nevertheless retain high thermal stability. Further, the mutant proteins were easily expressed recombinantly in *E. coli* in soluble form. On the other hand, the Tm1112 and ChBD mutants analyzed have significantly lower thermal stability (50 and 38 °C, respectively). The low thermal stability of Tm1112-BNP32 can be attributed to the deletion of multiple β-strands that likely results in significant disruption of the wild-type structure. Similar results are also seen in case of Sso7d, where deletion of the C-terminal α helix leads to a decrease in melting temperature by 46 °C.[20] In case of ChBD-rIgG, loss of thermal stability is likely due to an inadvertent amino acid insertion in a β-strand during oligonucleotide synthesis; all mutants in the ChBD library carry this insertion.

Table 3. Analysis of Thermal Stability for Wild-type Scaffold Proteins (WT) and Mutants[a]

Wild-Type or Target-binding Mutant	T_m or $T_{1/2}$ (°C)
Tm1112-WT	99.4
Tm1112-BNP32	50 ± 4.3[b]
MIT-WT	116.9 (116.7, 117.2)
MIT-rIgG	74.2 (74.1, 74.3)
Sso7d-WT	98[c]
Sso7d-Lysozyme	92.7[c]
Sso7d-rIgGm	85.5 (85.5, 85.6)
Sso6901-WT	97.2 (97.1, 97.3)
Sso6901-BCP	84.5 (84.5, 84.6)
Sso6901-BNP32	76.2 (76.0, 76.4)
ChBD-WT	>85[d]
ChBD-rIgG	38 ± 1.7[b]

[a]The melting temperature (T_m) was determined by differential scanning calorimetry. The temperature corresponding to the mid-point of irreversible thermal denaturation was determined from experiments with yeast cell surface displayed proteins. Unless stated otherwise, all data reported are T_m values. The average value from duplicate measurements is reported for all proteins, with the exception of Tm1112-WT, where a single measurement was performed. Numbers in parentheses show T_m values obtained in each experiment. The $T_{1/2}$ values reported are the mean from three separate experiments ± standard error of the mean. Also see Supplementary Figure S2. [b]$T_{1/2}$ value. [c]Values for Sso7d-WT[24] and Sso7d-Lysozyme[4] have been previously reported. [d]The wild-type ChBD protein has been reported to retain its secondary structure at 85 °C.[25]

CONCLUSIONS

A key limitation of combinatorial approaches is that any library screening method samples only a very small fraction of the sequence space generated by randomizing a scaffold protein. Here we have shown that, despite this limitation, a modestly sized superlibrary of hyperthermophilic scaffolds can yield binders to a diverse set of model targets. Interestingly, binding affinities of mutants isolated from the superlibrary are comparable to or higher than those obtained from an Sso7d library with 1000-fold higher diversity. Clearly, sequence diversity is important; it is intuitive that a scaffold library with very low diversity will be limited in its capability to generate binders. However, on the evidence of our data, it appears that the benefits of a modest increase in overall library size through scaffold diversification outweigh the gains from a large increase in the sequence diversity of a single scaffold library. To the best of our knowledge, this is also the first demonstration of the use of Sso6901, MIT, Tm1112, and ChBD as scaffolds for engineering molecular recognition, adding to the palate of alternate scaffolds available for generating binding proteins.

METHODS

Construction of Yeast Surface Display Libraries. The plasmid vector for yeast surface display (pCTCON) and the yeast strain EBY100 were kind gifts from Prof. K. Dane Wittrup (Massachusetts Institute of Technology, Cambridge, MA). A previously described Sso7d yeast library was used in this study.[4] Yeast display libraries for all other scaffolds and the Sso7d mRNA display library were constructed using previously published protocols[5,21] (see Supplementary Methods for details). Individual yeast display libraries were pooled together to obtain the superlibrary.

Library Screening. The yeast superlibrary was screened using a magnetic selection step followed by multiple rounds of FACS, as described.[5] The mRNA display library was incubated with biotinylated targets immobilized on Dynal biotin binder magnetic beads (Invitrogen, Carlsbad, CA). Mutants that bound to the beads were isolated, amplified by PCR, and subsequently transformed into yeast. This pool of mutants, in yeast display format, was further screened using FACS, and 6–10 clones from the pool of binders after the final round of FACS were sequenced. Detailed protocols can be found in Supplementary Methods.

Estimation of K_D. The apparent K_D was estimated using yeast cell surface titrations, for all mutants except rIgG binders, as described.[5] For MIT-rIgG and Sso7d-rIgGm, K_D was estimated using ELISA[22,23] (see Supplementary Methods for details). Errors associated with estimates of K_D are reported as 68% confidence intervals; these were calculated as described.[5]

Biophysical Characterization. Wild-type proteins and mutants were recombinantly expressed in *E. coli* with a C-terminal (all Sso7d mutants) or N-terminal (wild-type and mutant proteins from MIT, Sso6901, and Tm1112) 6xHistidine (6xHis) tag. All proteins were purified in a single step using a nickel column as described.[4] Detailed protocols can be found in Supplementary Methods. Melting temperatures of proteins dialyzed in PBS (wild-type Sso6901, wild-type MIT, Sso6901-BNP32, and Sso6901-BCP) or 50 mM sodium acetate buffer, pH 5 (wild-type Tm1112, MIT-rIgG, Sso7d-rIgG) were determined using a Nano DSC II differential scanning calorimeter (TA Instruments, Newcastle, DE). The thermal denaturation of ChBD-rIgG and Tm1112-BNP32 was studied using yeast cell surface experiments, previously described.[17,19] Protein samples in 20 mM sodium phosphate buffer (pH 7.4) were used for CD spectrocopy on a JASCO-815 spectropolarimeter (Jasco Inc., Easton, MD). CD spectra were recorded over 210–250 nm using a 50 nm/min scan rate, 0.1 nm pitch, 1 nm bandwidth, and 2 s response time.

■ ASSOCIATED CONTENT

Supporting Information

Detailed methods and supplementary figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

* Phone: 919-513-0129. Fax: 919-515-3465. E-mail: bmrao@ncsu.edu.

Author Contributions

†These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Rihe Liu and Steve Cotten (University of North Carolina, Chapel Hill) for help with mRNA display protocols. We also thank Paige Luck from the Department of Food Science, North Carolina State University (NCSU) and Dr. John van Zanten from the Biomanufacturing Training and Education Center (BTEC), NCSU for help with differential scanning calorimetry and CD experiments, respectively. This work was supported by the Defense Threat Reduction Agency (grant HDTRA1-10-1-0024).

■ REFERENCES

(1) Binz, H. K., Amstutz, P., and Pluckthun, A. (2005) Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol. 23*, 1257–1268.

(2) Lofblom, J., Frejd, F. Y., and Stahl, S. (2011) Non-immunoglobulin based protein scaffolds. *Curr. Opin. Biotechnol. 22*, 843–848.

(3) Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006) Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5869–5874.

(4) Gera, N., Hussain, M., Wright, R. C., and Rao, B. M. (2011) Highly stable binding proteins derived from the hyperthermophilic sso7d scaffold. *J. Mol. Biol. 409*, 601–616.

(5) Gera, N., Hussain, M., and Rao, B. M. (2012) Protein selection using yeast surface display. *Methods*, DOI: 10.1016/j.ymeth.2012.03.014.

(6) Virnekas, B., Ge, L., Pluckthun, A., Schneider, K. C., Wellnhofer, G., and Moroney, S. E. (1994) Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Res. 22*, 5600–5607.

(7) Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M., and Wittrup, K. D. (1999) Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J. Mol. Biol. 292*, 949–956.

(8) Park, S., Xu, Y., Stowell, X. F., Gai, F., Saven, J. G., and Boder, E. T. (2006) Limitations of yeast surface display in engineering proteins of high thermostability. *Protein Eng., Des. Sel. 19*, 211–217.

(9) Koide, A., Wojcik, J., Gilbreth, R. N., Hoey, R. J., and Koide, S. (2012) Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J. Mol. Biol. 415*, 393–405.

(10) Lipovsek, D., Lippow, S. M., Hackel, B. J., Gregson, M. W., Cheng, P., Kapila, A., and Wittrup, K. D. (2007) Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies. *J. Mol. Biol. 368*, 1024–1041.

(11) Watanabe, H., Matsumaru, H., Ooishi, A., Feng, Y., Odahara, T., Suto, K., and Honda, S. (2009) Optimizing pH response of affinity between protein G and IgG Fc: how electrostatic modulations affect protein-protein interactions. *J. Biol. Chem. 284*, 12373–12383.

(12) Takahashi, T. T., and Roberts, R. W. (2009) In vitro selection of protein and peptide libraries using mRNA display. *Methods Mol. Biol. 535*, 293–314.

(13) Barrios, Y., Jirholt, P., and Ohlin, M. (2004) Length of the antibody heavy chain complementarity determining region 3 as a specificity-determining factor. *J. Mol. Recognit. 17*, 332–338.

(14) Lamminmaki, U., Pauperio, S., Westerlund-Karlsson, A., Karvinen, J., Virtanen, P. L., Lovgren, T., and Saviranta, P. (1999) Expanding the conformational diversity by random insertions to CDRH2 results in improved anti-estradiol antibodies. *J. Mol. Biol. 291*, 589–602.

(15) Lee, C. V., Liang, W. C., Dennis, M. S., Eigenbrot, C., Sidhu, S. S., and Fuh, G. (2004) High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. *J. Mol. Biol. 340*, 1073–1093.

(16) Rock, E. P., Sibbald, P. R., Davis, M. M., and Chien, Y. H. (1994) CDR3 length in antigen-specific immune receptors. *J. Exp. Med. 179*, 323–328.

(17) Hackel, B. J., Kapila, A., and Wittrup, K. D. (2008) Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling. *J. Mol. Biol. 381*, 1238–1252.

(18) Koide, A., Gilbreth, R. N., Esaki, K., Tereshko, V., and Koide, S. (2007) High-affinity single-domain binding proteins with a binary-code interface. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6632–6637.

(19) Orr, B. A., Carr, L. M., Wittrup, K. D., Roy, E. J., and Kranz, D. M. (2003) Rapid method for measuring ScFv thermal stability by yeast surface display. *Biotechnol. Prog. 19*, 631–638.

(20) Shehi, E., Granata, V., Del Vecchio, P., Barone, G., Fusi, P., Tortora, P., and Graziano, G. (2003) Thermal stability and DNA binding activity of a variant form of the Sso7d protein from the archeon Sulfolobus solfataricus truncated at leucine 54. *Biochemistry 42*, 8362−8368.
(21) Huang, B. C., and Liu, R. (2007) Comparison of mRNA-display-based selections using synthetic peptide and natural protein libraries. *Biochemistry 46*, 10102−10112.
(22) Findlay, J. W., and Dillard, R. F. (2007) Appropriate calibration curve fitting in ligand binding assays. *AAPS J. 9*, E260−267.
(23) Lee, C. H., Park, K. J., Sung, E. S., Kim, A., Choi, J. D., Kim, J. S., Kim, S. H., Kwon, M. H., and Kim, Y. S. (2011) Engineering of a human kringle domain into agonistic and antagonistic binding proteins functioning in vitro and in vivo. *Proc. Natl. Acad. Sci. U.S.A. 107*, 9567−9571.
(24) Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I., and Ladenstein, R. (1998) Thermal unfolding of small proteins with SH3 domain folding pattern. *Proteins 31*, 309−319.
(25) Nakamura, T., Mine, S., Hagihara, Y., Ishikawa, K., Ikegami, T., and Uegaki, K. (2008) Tertiary structure and carbohydrate recognition by the chitin-binding domain of a hyperthermophilic Chitinase from Pyrococcus furiosus. *J. Mol. Biol. 381*, 670−680.