

An improved Protein G with higher affinity for human/rabbit IgG Fc domains exploiting a computationally designed polar network

Ramesh K.Jha[†], Tiziano Gaiotto[†],
Andrew R.M.Bradbury¹ and Charlie E.M.Strauss¹

Bioscience Division, MS M888, Los Alamos National Laboratory,
Los Alamos, NM 87545, USA

[†]To whom correspondence should be addressed.
E-mail: amb@lanl.gov (A.R.M.B.); cems@lanl.gov (C.E.M.S.)

Received February 4, 2014; revised February 4, 2014;
accepted February 10, 2014

Edited by Valerie Daggett

Protein G is an IgG binding protein that has been widely exploited for biotechnological purposes. Rosetta protein modeling identified a set of favorable polar mutations in Protein G, at its binding interface with the Fc domain of Immunoglobulin G, that were predicted to increase the stability and tighten the binding relative to native Protein G, with only a minor perturbation of the binding mode seen in the crystal structure. This triple mutant was synthesized and evaluated experimentally. Relative to the native protein G, the mutant showed a 3.5-fold enhancement in display level on the surface of yeast and a 5-fold tighter molar affinity for rabbit and human IgG. We attribute the improved affinity to a network of hydrogen bonds exploiting specific polar groups on human and rabbit Fc. The relative specificity increased as well since there was little affinity enhancement for goat and mouse Fc, while the affinity for rat Fc was poorer by half. This designed Protein G will be useful in biotechnological applications as a recombinant protein, where its improved affinity, display and specificity will increase antibody capture sensitivity and capacity. Furthermore, the display of this protein on the surface of yeast introduces the concept of the use of yeast as an affinity matrix.

Keywords: Computational protein design/Fc domain/
Immunoglobulin G/Polar interaction/Protein G

Introduction

Protein G (PrG) is an immunoglobulin G (IgG) binding protein found on the surface of group G streptococci (Fahnestock *et al.*, 1986). It is thought to be involved in pathogenesis, facilitating evasion of the immune system by binding host IgG and thereby camouflaging the bacterial surface with host proteins (Goward *et al.*, 1993) and/or by inhibiting signal transduction by masking the Fc from the immune system. PrG consists of three Fc-binding domains (C1, C2 and C3), as well as an albumin binding portion (Akerstrom *et al.*, 1987). The Fc domain is common to all antibodies and plays a role in

antibody effector function, signaling and antibody fate. The crystal structure of the C2 domain bound to an Fc domain (Sauer-Eriksson *et al.*, 1995) shows that it involves predominantly polar interactions across the interface with multiple hydrogen bonds and salt bridges. Indeed, there are only three significant hydrophobic interactions: burial of two methionine residues and one isoleucine on the Fc domain side. In comparison, the interaction between protein A and Fc is mediated predominantly by hydrophobic interactions (Deisenhofer, 1981).

The Fc-binding property of PrG has led to numerous biotechnological applications of PrG as an affinity reagent. These include antibody purification or detection (Stahl *et al.*, 1993), siRNA delivery (Bae *et al.*, 2013) and immuno-rolling circle amplification (Akter *et al.*, 2011). Higher affinity variants of PrG's Fc-binding domain would significantly enhance the biotechnological utility of this potent reagent. We sought to improve the interaction between PrG and the Fc domains of rabbit and human IgG antibodies by using Rosetta (Rohl *et al.*, 2004; Leaver-Fay *et al.*, 2011) to identify interface amino acids that could improve affinity if mutated, and then experimentally examining the effects of these mutations on Fc binding.

Design of novel protein–protein interactions is a current challenge that paves the way for advanced therapeutics, affinity reagents and self-assembling systems. Recent years have seen advances in computational redesign of protein–protein interfaces as well as the *de novo* design of binding partners (Fleishman *et al.*, 2011c). However, protein interface design efforts often fail, and inaccuracy in modeling buried polar interactions appears to be a major failure mode (Der *et al.*, 2013; Stranges and Kuhlman, 2013). Instead of enhanced polar interaction, most successful examples of protein interface redesign have relied upon steric and hydrophobic interactions; including a hydrophobic cleft (Jha *et al.*, 2010), a designed hydrophobic motif (Karanicolas *et al.*, 2011), a designed hydrophobic hotspot (Fleishman *et al.*, 2011b), mutating polar to non-polar amino acids, or extending the side chain in non-polar amino acids (Liu *et al.*, 2007; Sammond *et al.*, 2007; Fridman *et al.*, 2013). Examples of *de novo* designs exploiting polar mutations are rare. Polar interactions have resulted in antibody-affinity improvement (Lippow *et al.*, 2007) and in some cases have been used to confer specificity (Havranek and Harbury, 2003). In two other recent notable successes, the polar mutants emerged from random directed evolution techniques not by blind computational determination (Fleishman *et al.*, 2011b; Karanicolas *et al.*, 2011). Polar interactions are attractive because of the potential for a large gain in both binding energy and specificity from very few mutations (Joachimiak *et al.*, 2006; Stranges *et al.*, 2011; Der *et al.*, 2012; Procko *et al.*, 2013). Here, we present a successful polar residue redesign strategy that improves an important, ubiquitous, biotechnology reagent.

[†]These two authors contributed equally to this work.

Most reports involving antibodies and improved affinity are concerned with finding a better affinity antibody for a specific target. Instead, here we are modifying PrG, which interacts with the Fc domain of the antibody, well away from the variable region that is responsible for antigen binding. Thus, our recombinant PrG is aimed at enhancing a general-purpose affinity reagent for human/rabbit antibodies.

Our redesign problem is one-sided: we are engineering PrG while leaving its target human IgG-Fc protein in its unmodified native form. This restriction increases the difficulty as it denies convenient strategies possible in two-sided design, such as swapping a polar pair across the interface, or entirely erasing troublesome polar interactions. Since the prospect of a redesign further improving upon a natural, highly evolved, tight binder (sub micro molar) is challenging for *in silico* design, we are pleased to report success here.

Materials and methods

Refining the structure of the complex

The crystal structure (PDB code 1FCC) (Sauer-Eriksson *et al.*, 1995) is a relatively low resolution structure (3.20 Å) containing the bound form of human Fc (hFc; Chain A) and PrG (Chain C). The structure of the complex was refined using the ‘constrained relax’ protocol of the Rosetta macromolecular modeling suite to prepare a template free of clashes in the Rosetta potential (Nivon *et al.*, 2013); this was the starting point for all subsequent *in silico* mutagenesis. Ten relaxation trajectories were performed and the lowest energy structure was then used for the sequence redesign.

Redesign of the interface sequence

We used the protein interface design protocol (Fleishman *et al.*, 2011a) in release 3.4 of the Rosetta macromolecular modeling suite (Bonneau *et al.*, 2001; Rohl *et al.*, 2004; Leaver-Fay *et al.*, 2011) to predict mutations that can improve binding affinity. ‘Interface residues’ were defined as any pair of residues between the hFc and PrG whose side chains’ centroids were within 8 Å. The interface redesign protocol cycles through four steps. In the first step, PrG is further refined with small backbone perturbations and side chain rotamer variations focused on the interface region, along with rigid body minimization of the two domains. In the second step, the interface sequence is modified and the interface repacked by side chain rotamer sampling and rigid body minimization of the total Rosetta energy (Gray *et al.*, 2003). During the first two steps the energy function is softened to tolerate small clashes in the repulsive Lennard-Jones potential term (‘soft-rep’ protocol), which greedily accelerates sequence changes. The third step repeats the second step with a repulsive Lennard-Jones potential term to anneal the residual atomic clashes. In the final step, small torsion angle changes to the backbone are allowed as well. Ten cycles of steps 2–4 are applied to generate each model. Around 1900 models were generated.

Estimation of binding energy from computation

Binding energies are computed from the difference of the Rosetta energy in the bound and unbound forms. An implicit assumption is that the backbone structure of the unbound domain is unchanged from the bound form though the rotamers may vary. While that assumption could be problematic

for proteins with large allosteric rearrangements upon binding, it is appropriate here where the feature of interest is the difference in binding energy between the native and the near native designed interactions.

Selection of mutations for experiments

We screened the design ensemble, keeping the 10% of the 1900 models with the lowest total Rosetta energy. We further filtered these high stability complexes by binding energy, keeping 30 with the tightest binding energy. Models were inspected by eye using the PyMOL molecular graphics tool (Schrödinger, LLC) and, as described below, we further narrowed our focus to five mutation sites. The sequence diversity in the 30 models at these positions is shown in Fig. 1.

Comparison with the native sequence

As a basis for comparison, we followed the same protocols above without letting the sequence vary from the native. This restriction reduces the search space, so for this we generated 30 models. As mentioned above, we chose the 10% of the 30 models by the total Rosetta energy. Figure 2B shows the model with the lowest binding energy out of top three models based on total Rosetta energy.

Comparative modeling of goat, mouse and rat IgG Fc domain

Using the PrG and hFc as a template, we substituted the amino acids on Fc, which were within 5 Å of the PrG chain with the corresponding amino acids of goat, mouse and rat (based on multiple sequence alignment, Supplementary Fig. S1). For each fixed sequence, 30 trajectories were carried out by using both the native and mutated PrG sequences. The binding energy reported (Supplementary Table SI) is the average of the top three models based on total Rosetta energy for each subtype. Mouse and rat had identical interface substitutions and thus are indistinguishable at the level of comparative modeling.

Random protein–protein docking to estimate the energy funnel

On the hFc interface, 10 200 protein–protein docking trajectories (Wang *et al.*, 2007; Chaudhury *et al.*, 2011) were sampled with both the native and mutated PrG. Each docking trajectory was initialized with a random rigid body translational and rotational perturbation of the PrG domain over a standard deviation of 3 Å and 8°, respectively. After rotamer relaxation, the total energy and binding energy were tabulated (Fig. 3).

Yeast surface display of PrG and ERH mutant

The C2 binding domain gene from Streptococcal PrG (PrG-WT) was synthesized by Genscript with the codon usage optimized for yeast expression (*Saccharomyces cerevisiae* strain). A secretion tag to promote display was placed at the N-terminal and the strand 11 tag from split green fluorescent protein (GFP) (Cabantous *et al.*, 2005) was added to the C-terminal to allow quantitative normalization of the display of full-length PrG. The mutant ERH (PrG-ERH—named for the three altered amino acids) was generated using two oligos and the Phusion site-directed mutagenesis kit (Thermo Scientific). The two genes were subcloned into pDNL6-GFP11 (Ferrara *et al.*, 2011), transformed into EBY100 yeast strain following the Yeast transformation kit protocol (Sigma) and verified by sequencing. Growing conditions and strands



Fig. 1. Position-specific sequence diversity in the ensemble of models. The histogram (LOGO) shows the sequence recurrence frequency at five positions based on the best 1.5% of 1900 design models for PrG. There is a distinct preference for mutation of two residues to highly polar residues (A24E and K28R). Our experimental structure also included the less abundant mutation V29H. Each column header indicates the native PrG sequence.

1–10 GFP complementation assays were performed as previously described (Ferrara *et al.*, 2011).

Measurement of binding affinity

The molecular affinities (dissociation constant or K_d) of wild-type (PrG-WT) and mutant (PrG-ERH) towards IgG were calculated directly on the yeast surface by fitting the display-normalized fluorescence of serial titrations of fluorescently labeled antibody (Chao *et al.*, 2006) to yeast cells displaying the different PrG forms. This method of calculating affinity (Boder and Wittrup, 2000) shows consistency with other techniques (Gai and Wittrup, 2007). The AlexaFluor 647-conjugated IgGs, purchased from Jackson Immuno Research, were determined to have no fluorescent crosstalk with the GFP, allowing two-channel flow cytometer assays to separately measure display and antibody capture. Figure 4D–F show un-normalized data and the supplementary material contains normalized curves. This process was repeated for five Fc species: human, rabbit, mouse, rat and goat.

Results

Interface redesign of PrG and hFc

The crystal structure of the bound form of PrG and hFc was our starting template structure (PDB code: 1FCC) (Sauer-Eriksson *et al.*, 1995). Since it was low resolution (3.2 Å), and, generally, crystal structures are not an exact local minimum of the Rosetta potential, we refined this to relieve artifactual clashes that would mask the subsequent smaller improvements arising from *in silico* mutations. The lowest energy refined native deviated only 0.15 Å root mean square deviation (RMSD) from the crystal structure backbone.

As detailed in the Materials and methods section, we redesigned the PrG sequence at its interface with Fc. This

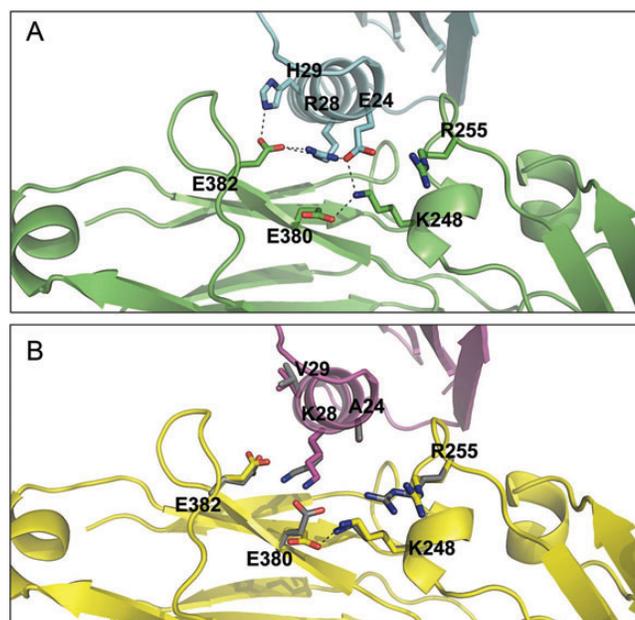


Fig. 2. Predicted structure of the PrG and hFc interface. (A) A Rosetta predicted model of the hFc (green) bound to the designed ERH mutant of PrG (cyan) shows a partially buried, polar network forming connections across the interface. (B) The Rosetta-relaxed structure of native-sequence PrG (magenta) and hFc (yellow). The overlaid (3.2 Å resolution) crystal structure before refinement (gray) shows the Fc polar groups along the rim of the interface have fewer interactions to the PrG than the mutant.

produced an ensemble of 1900 model structures, each of which is a local minimum of the Rosetta energy found by a Monte Carlo search of sequence and structure. From these we chose the 30 best models that had improved both the binding affinity and overall stability. Intriguingly, the binding mode of the designed structures had shifted by 1.5 Å (RMSD) from the crystal structure binding mode; this predicted shift may be what enables the new hydrogen bonding interactions we describe below.

Selection of one sequence for experimental analysis

At this point, the small spread in energies among the top models was not usefully discriminative as a sole criterion for selecting a minimal set of mutations. We examined the overall position-specific substitution frequencies observed across this ensemble of low energy models: 19 residue positions had designed mutations. The most intriguing consensus was a consecutive epitope of six amino acids from position 24 to 29 on PrG, which showed a very strong propensity for polar amino acids (Fig. 1).

From these six sites we chose a triple mutant for experimental study as follows. As seen in the Logo plot (Crooks *et al.*, 2004) (Fig. 1), E27 had a strong propensity for its native residue, so conservative parsimony dictated leaving it unchanged in our experimental sequence. Residue T25 had its propensity dispersed across many amino acid types, so we avoided experimental mutation at this ambiguous site. This is also in accordance to the observation made earlier with respect to PrG mutations, where E27 was shown to be important for binding while T25 was tolerant to mutations (Sloan and Hellinga, 1998). Position 26 was also omitted because the buried side chain is directed away from the interface.

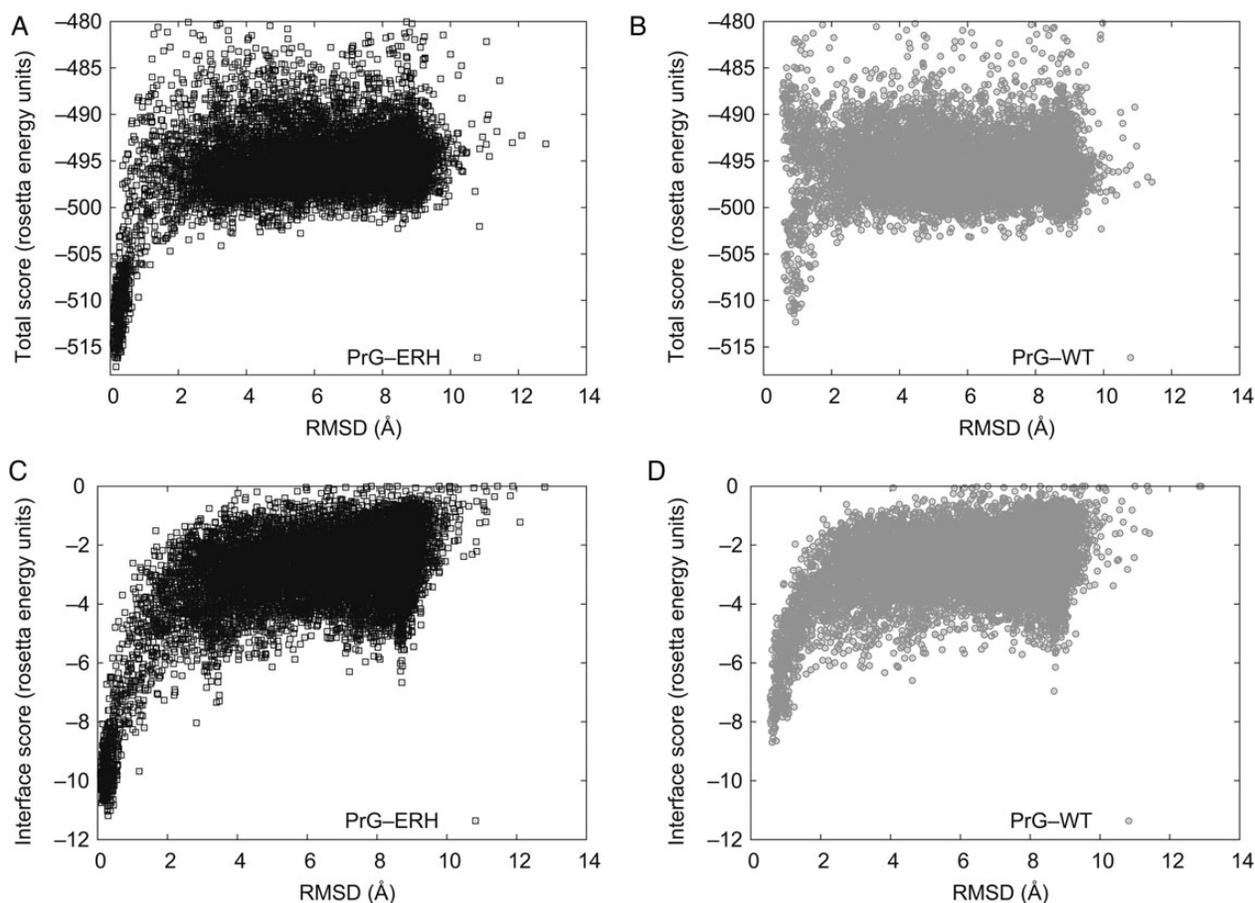


Fig. 3. Scatter plot of predicted model energy versus deviation from the designed structure for a series of docking perturbations. Each point on the plot is a Rosetta model perturbed from the selected minimum energy conformation for mutant complex or relaxed native complex and their backbone deviation referenced to the selected mutant complex with lowest binding energy. Both the triple mutant ERH PrG (black/left) and the relaxed native (gray/right) show the onset of a strong correlation between model's energy and its atomic deviation from the structure of our designed model. This 'funnel' is deeper, sustained over a wider range, and more populated for the mutant than the native. For expository convenience, the native models are shown referenced to the designed structure not their relaxed native structure: none of the low energy native models are within ~ 1 Å of the designed mutant model (RMS C-alpha superposition) so their funnel minimum is displaced. (A and B): The total energy indicates the relative stability of the complex. (C and D): The interface score is the bound minus unbound total energy.

This left us with three positions which showed mutations from the non-polar to polar amino acids (A24, V29) or from the polar to more polar amino acids (K28). The most intriguing substitution was K28R. Our designs remodeled the existing inter-domain K28-E380 interaction to a more robust hydrogen bonding network (Fig. 2A). The weakness of this original interaction in the crystal structure is inferred from the Rosetta potential's tendency to flip the E380 rotamer to a different, intra-domain, interaction E380-K248 (Fig. 2B). Thus, even though this position had an imperfect consensus in the models (Fig. 1), the prospect of forming this enhanced hydrogen bonding network was attractive, hence we selected K28R.

The Logo plot (Fig. 1) shows only a weak propensity for our selected mutant V29H. However, the T and K model mutations at position 29 were involved in intra-domain hydrogen bonding which stabilized PrG but did not contribute to the binding energy to the Fc domain. Hence, we chose the V29H mutation over the T and K mutations, despite their higher model propensity, because it showed an inter-domain interaction. Moreover, in models that contained our other four selections (A24E, T25, E27 and K28R), it balanced the R28-E382 hydrogen bond. In the native crystal structure position, V29 appeared to have an unfilled void at the interface.

Models containing polar mutations A24E as well as K28R and V29H, dubbed PrG-ERH, showed formation of a hydrogen bond network in some of the models (Fig. 2A). In both the crystal structure and in the relaxed native complex, this strongly defined hydrogen bond network appeared to be absent (Fig. 2B). The inter-domain network offered the prospect of large energetic gain over the native in both the stability and affinity (Fig. 3).

Evidence for a strong basin of attraction around the lowest energy model

Experience has shown that a hallmark of a well-predicted protein-protein interaction is a deep energy well in the vicinity of the lowest energy conformation (Smith and Sternberg, 2002; Gray et al., 2003; London and Schueler-Furman, 2008). This may be assessed by random perturbations of the docking of the domains, and graphically exhibited by a scatterplot of each perturbed model's energy versus root mean square (RMS) C-alpha distance from the reference structure. Characteristically, after thousands of docking trajectories, one observes a large cloud of points (models) at high energy and large deviations that show no correlation to energy. The signature of a good interface is a subset of models forming an energy 'funnel' that descends

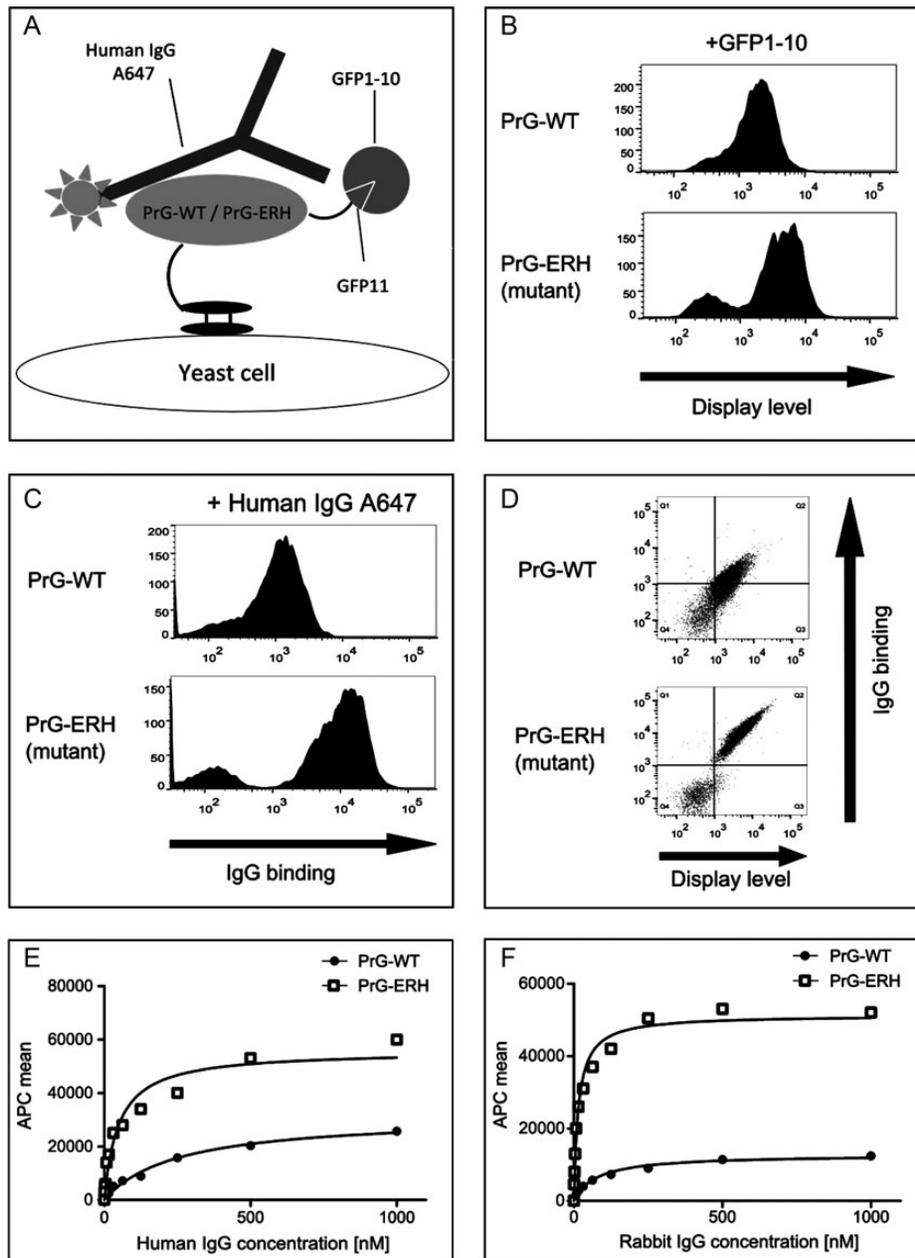


Fig. 4. Experimental yeast display of the native and mutant PrG measures affinity and display levels by flow cytometry. (A) PrG displayed on the surface of yeast binds the Fc stalk region of a fluorescently labeled IgG antibody. The PrG C-terminal strand 11 tag complements with exogenous split GFP providing an independent channel to monitor display. (B) Histogram of cell-to-cell variation in GFP fluorescence shows mutant ERH PrG has higher display level than native (WT). (C) Histogram of antibody fluorescence tag shows increased capture by mutant relative to WT. (D) The two-axis plot shows each cell's simultaneous antibody tag fluorescence versus the displayed GFP tag fluorescence. The population of cells both displaying PrG (or the mutant) and binding IgG are gated to compute the mean affinity per molecule after normalization for display. (E) and (F) Titration series of the (un-normalized) mean antibody fluorescence as a function of IgG concentration for human (E) and rabbit (F) shown for both mutant (ERH) and native (WT) PrGs. See Table I for a summary for affinity changes for five species.

from that larger cloud to much lower energy, converging close to the designed structure at the lowest energy. Also, anticipated is an absence of additional low energy funnels representing competing interface geometries. The deeper and more populated the basin of attraction, the more confident we are that our design structure will be the low energy experimental conformation.

The scatter plot (Fig. 3) of 10 200 randomly docked models shows the binding energy funnel of the native (gray circles) and redesigned A24E, K28R, V29H mutant PrG (black squares). Since they share similar overall interfaces, both native and redesigned PrG show qualitatively similar funnel

shapes. However, the redesigned PrG (PrG-ERH) shows a quantitatively deeper, broader and more densely populated funnel than the native PrG (PrG-WT), which we attribute to the large energy gain of the new hydrogen bonding network.

The basin of attraction of the native deviated ~ 1 Å from the best binding model of our mutant design. While small, the shift is wider than the width of the funnel, making it significant. This indicates a rigid body twist between the binding partners driven by these strong polar interface groups. Despite the shift, the mutant's funnel itself was deeper and more heavily sampled than the native, in both the scatter plots for

Table I. Measured molar binding constant for mutant (ERH) and native (WT) PrG against five IgG species.

	K_d PrG-WT, (nM)	K_d PrG-ERH, (nM)	Affinity gain	Correlation coefficient
Rabbit	77	16	5x	0.98
Human	250	49	5x	0.94
Goat	320	300	1x	0.97
Rat	670	1300	0.5x	0.97
Mouse	920	640	1.5x	0.99

total energy and binding energy. This dominating basin of attraction increases our confidence in the model precision.

Our Rosetta relaxed structures of both the native sequence and redesigned protein sequence showed a docked complex that deviated from the crystal structure by 1.5 Å (Supplementary Fig. S2). Essentially all of this deviation arises in a rigid body shift of the inter-domain binding mode, not in our small changes to the interface backbone and side chains. While surprisingly large, this deviation is within the 3.2 Å resolution of the 1FCC crystal structure. It has been seen that Rosetta relaxation can often improve (refine) the accuracy of a low resolution cryoEM density maps, nuclear magnetic resonance or crystal structure when the starting point is within a few angstroms of being correct (Qian *et al.*, 2007; DiMaio *et al.*, 2009). However, we must cautiously note that while we are optimistic that we have refined the model to higher accuracy at the interface, it is possible that 1.5 Å might exceed the convergence zone in which Rosetta can globally refine the X-ray structure of a large complex.

Experimental measurements

In order to assess the Rosetta predictions described above, separate genes for the native and mutant PrG were synthesized using yeast-preferred codons and cloned into a yeast display vector (Ferrara *et al.*, 2012). This allows PrG to be displayed on the surface where it is able to bind exogenous, fluorescently labeled antibody (Fig. 4A). The displayed PrG also contained at its C-terminus a short GFP strand 11 (S11) tag. This can be complemented with exogenously added GFP strands 1–10, which is non-fluorescent until it binds to S11 (Cabantous *et al.*, 2005). S11 has been engineered to avoid disruption of expression, folding, export or display in the yeast, and has been previously used to normalize protein display on yeast (Ferrara *et al.*, 2011). The asymptotic affinity (K_d) is calculated from a titration series of antibody concentrations normalized to the concentration of the displayed PrG, assuring that the increased IgG binding activity obtained on yeast cells is not an artifact of the higher PrG ERH display (Colby *et al.*, 2004) (Supplementary Fig. S3).

The GFP fluorescence histogram of cell-to-cell variation shown in Fig. 4B indicates that the per-cell display of mutant PrG (bottom panel) is increased by about 3.5-fold relative to the native (top panel). Figure 4D shows that the fluorescently labeled antibody captured by PrG varies linearly with the display levels and thus there is negligible non-specific binding. Table I summarizes the molar affinities (asymptotic binding constant K_d) derived from fits to the mean fluorescence in an antibody titration series (examples shown in Fig. 4E and F for human and rabbit IgGs).

The mutations also increased specificity to human IgG. Table I presents the binding affinity of the native and mutant PrG variants for IgGs from five species: rabbit, human, rat, mouse and goat. The ERH mutant shows 5-fold higher binding affinity than the native protein against human and rabbit antibodies, but little or no improvement for mouse, rat and goat IgGs.

There are no published structures of PrG complexed with goat, rat, mouse or rabbit Fc. Nevertheless, we endeavored to model the relative affinity of the new ERH mutant to goat and rat Fc by naively threading their unique interface residues onto the hFc structure. The predicted binding energies for these were negligibly different from hFc, which is contrary to the large experimentally measured affinity differences (Supplementary Table SI). To explain this, we note that while the human and rabbit share identical Fc interface residues, their measured K_d is quite different, indicating that strong differences can arise from residues outside the modeled interface region. Thus, comparative modeling restricted to just the interface is invalidated for these unknown structures.

Discussion

Our methodology relied on a computational Monte Carlo search across the structure and sequence of about 20 residue positions in PrG to optimize the Rosetta all-atom energy (Rohl *et al.*, 2004). Because the Rosetta design process does not use molecular dynamics, explicit waters or exhaustive enumeration, it can efficiently explore a large combinatorial space of structure and sequence with relatively modest computing effort.

A dual criteria of improving both the complex stability and the binding affinity was imposed on the model selection. We parsimoniously down-selected to a set of three PrG mutations in close proximity (A24E, K28R and V29H), which were predicted to form a highly interconnected hydrogen bond network in the Rosetta generated *de novo* computational models.

Importantly, this was a *blind* design, based solely on the potential energy model, without directly imposing information about naturally occurring mutations at these sites. It was therefore reassuring to see that the same glutamate mutation (24E) has evolved in many natural variants of PrG. A Psiblast alignment (Altschul *et al.*, 1997) finds ~30% of the closest 100 sequence matches to PrG contain A24E; a few percent contain K28R, but none contain V29H.

Polar mutations can provide large gains in binding (Fleishman *et al.*, 2011b; Karanicolas *et al.*, 2011). However, the literature indicates that *in silico* design including a polar substitution is hazardous (Stranges and Kuhlman, 2013). The key impediment is that each polar residue adds a desolvation cost against binding that must be compensated with a sufficiently good interfacial interaction. Accurate calculation of both of these large energy terms is currently difficult often making their small difference, the predicted binding energy, unreliable even in sign. For a buried polar interaction, the solvation penalty is unforgiving, since it is paid regardless of whether a compensating polar interaction was well designed or not.

Our success with a triple polar mutation suggests an insight to overcome this hurdle: *target long flexible polar groups at the rim of the binding interface*. Our utilitarian definition of ‘rim’ is any residue buried by cross-interface contacts that

could, after plausible side chain rearrangement, stretch its polar groups into a solvated region. Our polar mutations in PrG were along the interface rim. If the design turns out to be accurate, then one gets the full enthalpy benefit of the satisfied interfacial polar interactions, but otherwise, if a designed polar interaction fails, the rim location may allow flexible polar residues to rotate into contact with water molecules. Our hypothesis is that this rescues it from the desolvation penalty and thereby tolerates an imperfect design.

This ‘rim’ strategy offers the designer latitude to risk more polar substitutions. For example, we were concerned that the V29H mutation could be deleterious; in some Rosetta models histidine is seen in alternate conformation and not hydrogen bonding with E382. However, being at the rim of the interface, potential rotamer variations would allow the residue to solvate, potentially tolerating this hydrophobic-to-polar mutation. Indeed, the native complex appears to use this strategy as well: the native Fc interface has polar residues uncompensated by the native PrG but these lie along the rim of the binding patch.

Why does the native PrG lack this optimization? Given its role in pathogenesis, it is likely that for Streptococci expressing PrG, promiscuity with respect to Fc subtypes would be preferable to specific high affinity binding to the Fc from any single species. Consequently, the sequence and structure would be expected to forgo optimization for the human/rabbit-Fc specific polar groups we targeted. The empirical selectivity improvement for human and rabbit Fc relative to other species is consistent with this hypothesis. Indeed, since some of our inferred mutations are found in sequence homologs, it may be that some of those homologs belong to specialists for the human/rabbit subtype Fc.

Experimental evaluation showed a 5-fold affinity improvement for binding to Fc from human and rabbit, achieving <16 nM K_d in rabbit IgG and <49 nM K_d human IgG. We also observed a 3.5-fold increase in yeast display, likely reflecting increased solubility from the polar surface. This results in a yeast antibody capture system with >17 -fold higher efficiency, and improved selectivity compared with native PrG. In addition to being useful as a replacement for PrG in many biotechnological applications, the display of this mutant PrG on the surface of yeast, also introduces the concept of renewable affinity matrices, akin to the use of Staphylococci to purify proteins based on the natural expression of protein A, but using improved recombinant proteins instead.

Conclusion

Polar interactions can help increase the specificity and sensitivity of protein–protein interactions and the present example succeeds in designing hydrogen bonds and salt bridges at the rim of an interface, yielding a novel PrG variant with improved IgG Fc-binding capacity.

Supplementary data

Supplementary data are available at *PEDS* online.

Funding

This work was supported by the Defense Threat Reduction Agency [CBCALL12-LS-6-0622 to C.E.M.S.], LANL Institutional Computing [W11_SynBio to C.E.M.S.] and

LANL Laboratory Directed Research and Development program [20120029DR to C.E.M.S.], and by the National Institute of Health [5U54DK093500-02 to A.R.M.B.]

References

- Akerstrom, B., Nielsen, E. and Bjorck, L. (1987) *J. Biol. Chem.*, **262**, 13388–13391.
- Akter, F., Mie, M. and Kobatake, E. (2011) *Anal. Biochem.*, **416**, 174–179.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) *Nucleic Acids Res.*, **25**, 3389–3402.
- Bae, J., Mie, M. and Kobatake, E. (2013) *Biotechnol. Lett.*, **35**, 2081–2089.
- Boder, E.T. and Wittrup, K.D. (2000) *Methods Enzymol.*, **328**, 430–444.
- Bonneau, R., Tsai, J., Ruczinski, I., Chivian, D., Rohl, C., Strauss, C.E.M. and Baker, D. (2001) *Proteins*, Suppl 5, 119–126.
- Cabantous, S., Terwilliger, T.C. and Waldo, G.S. (2005) *Nat. Biotechnol.*, **23**, 102–107.
- Chao, G., Lau, W.L., Hackel, B.J., Sazinsky, S.L., Lippow, S.M. and Wittrup, K.D. (2006) *Nat. Protoc.*, **1**, 755–768.
- Chaudhury, S., Berrondo, M., Weitzner, B.D., Muthu, P., Bergman, H. and Gray, J.J. (2011) *PLoS ONE*, **6**, e22477.
- Colby, D.W., Kellogg, B.A., Graff, C.P., Yeung, Y.A., Swers, J.S. and Wittrup, K.D. (2004) *Methods Enzymol.*, **388**, 348–358.
- Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) *Genome Res.*, **14**, 1188–1190.
- Deisenhofer, J. (1981) *Biochemistry*, **20**, 2361–2370.
- Der, B.S., Machius, M., Miley, M.J., Mills, J.L., Szyperski, T. and Kuhlman, B. (2012) *J. Am. Chem. Soc.*, **134**, 375–385.
- Der, B.S., Jha, R.K., Lewis, S.M., Thompson, P.M., Guntas, G. and Kuhlman, B. (2013) *Proteins*, **81**, 1245–1255.
- DiMaio, F., Tyka, M.D., Baker, M.L., Chiu, W. and Baker, D. (2009) *J. Mol. Biol.*, **392**, 181–190.
- Fahnestock, S.R., Alexander, P., Nagle, J. and Filpula, D. (1986) *J. Bacteriol.*, **167**, 870–880.
- Ferrara, F., Listwan, P., Waldo, G.S. and Bradbury, A.R. (2011) *PLoS ONE*, **6**, e25727.
- Ferrara, F., Naranjo, L.A., Kumar, S., Gaiotto, T., Mukundan, H., Swanson, B. and Bradbury, A.R. (2012) *PLoS ONE*, **7**, e49535.
- Fleishman, S.J., Leaver-Fay, A., Corn, J.E., et al. (2011a) *PLoS ONE*, **6**, e20161.
- Fleishman, S.J., Whitehead, T.A., Ekiert, D.C., Dreyfus, C., Corn, J.E., Strauch, E.M., Wilson, I.A. and Baker, D. (2011b) *Science*, **332**, 816–821.
- Fleishman, S.J., Whitehead, T.A., Strauch, E.M., et al. (2011c) *J. Mol. Biol.*, **414**, 289–302.
- Fridman, Y., Gur, E., Fleishman, S.J. and Aharoni, A. (2013) *Proteins*, **81**, 341–348.
- Gai, S.A. and Wittrup, K.D. (2007) *Curr. Opin. Struct. Biol.*, **17**, 467–473.
- Goward, C.R., Scawen, M.D., Murphy, J.P. and Atkinson, T. (1993) *Trends Biochem. Sci.*, **18**, 136–140.
- Gray, J.J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C.A. and Baker, D. (2003) *J. Mol. Biol.*, **331**, 281–299.
- Havranek, J.J. and Harbury, P.B. (2003) *Nat. Struct. Biol.*, **10**, 45–52.
- Jha, R.K., Leaver-Fay, A., Yin, S., Wu, Y., Butterfoss, G.L., Szyperski, T., Dokholyan, N.V. and Kuhlman, B. (2010) *J. Mol. Biol.*, **400**, 257–270.
- Joachimiak, L.A., Kortemme, T., Stoddard, B.L. and Baker, D. (2006) *J. Mol. Biol.*, **361**, 195–208.
- Karanicolas, J., Corn, J.E., Chen, I., et al. (2011) *Mol. Cell*, **42**, 250–260.
- Leaver-Fay, A., Tyka, M., Lewis, S.M., et al. (2011) *Methods Enzymol.*, **487**, 545–574.
- Lippow, S.M., Wittrup, K.D. and Tidor, B. (2007) *Nat. Biotechnol.*, **25**, 1171–1176.
- Liu, S., Zhu, X., Liang, H., Cao, A., Chang, Z. and Lai, L. (2007) *Proc. Natl Acad. Sci. USA*, **104**, 5330–5335.
- London, N. and Schueler-Furman, O. (2008) *Biochem. Soc. Trans.*, **36**, 1418–1421.
- Nivon, L.G., Moretti, R. and Baker, D. (2013) *PLoS ONE*, **8**, e59004.
- Procko, E., Hedman, R., Hamilton, K., et al. (2013) *J. Mol. Biol.*, **425**, 3563–3575.
- Qian, B., Raman, S., Das, R., Bradley, P., McCoy, A.J., Read, R.J. and Baker, D. (2007) *Nature*, **450**, 259–264.
- Rohl, C.A., Strauss, C.E., Misura, K.M. and Baker, D. (2004) *Methods Enzymol.*, **383**, 66–93.
- Sammond, D.W., Eletr, Z.M., Purbeck, C., Kimple, R.J., Siderovski, D.P. and Kuhlman, B. (2007) *J. Mol. Biol.*, **371**, 1392–1404.
- Sauer-Eriksson, A.E., Kleywegt, G.J., Uhlen, M. and Jones, T.A. (1995) *Structure*, **3**, 265–278.

- Sloan,D.J. and Hellinga,H.W. (1998) *Protein Eng.*, **11**, 819–823.
- Smith,G.R. and Sternberg,M.J. (2002) *Curr. Opin. Struct. Biol.*, **12**, 28–35.
- Stahl,S., Nygren,P.A., Sjolander,A. and Uhlen,M. (1993) *Curr. Opin. Immunol.*, **5**, 272–277.
- Stranges,P.B. and Kuhlman,B. (2013) *Protein Sci.*, **22**, 74–82.
- Stranges,P.B., Machius,M., Miley,M.J., Tripathy,A. and Kuhlman,B. (2011) *Proc. Natl Acad. Sci. USA*, **108**, 20562–20567.
- Wang,C., Bradley,P. and Baker,D. (2007) *J. Mol. Biol.*, **373**, 503–519.