Computational design of interleukin-2 mimetics

Shawn Yu

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee

David Baker, Chair

Valerie Daggett

Marion Pepper

Program Authorized to Offer Degree:

Department of Bioengineering

University of Washington

Abstract

Computational design of interleukin-2 mimetics

Shawn Yu

Chair of Supervisory Committee:
Professor David Baker
Biochemistry

Interleukin-2 is a cytokine that plays a central role in immune system homeostasis, exerting paradoxical immunostimulatory and immunoregulatory effects based on its interactions with various receptor subunits differentially expressed across different cell types. It has been explored for a wide range of potential therapeutic applications. For my work in this dissertation, I have computationally designed minimalist interleukin-2 mimetics that retain only the structural elements necessary for binding and signal transduction with decreased dependence on biochemical context. It is my hope that these computationally designed mimetics can form the basis of useful engineered therapeutics in applications where it is desirable to shift the balance of the immune response.

Acknowledgments

I would like to express my deep gratitude for those who have helped me on the work presented in my dissertation. First, I would like to thank the members of the protein interface design team for helpful discussions.

Sarel Fleishman and Eva-Maria Strauch developed and polished many of the computational tools used in interface design. I must also recognize David La, Gabriel Rocklin, Lucas Nivon, Neil King, Jacob Bale, David Kim, Ray Wang, and Gustav Oberdorfer for sharing valuable computational protocols that I used throughout various stages of my doctoral research.

James Moody, Aaron Chevalier, and Christopher Bahl have been incredibly patient and helpful in guiding me along as I transitioned into experimental work, but I have also relied extensively on the experience of Cassie Bryan, Matt Bick, Jeremy Mills, Justin Siegel, Possu Huang Jorgen Nelson, Jiayi Dou, Austin Day, Luke Dang, and many others.

Experimental assistance was provided by Lauren Carter and the rest of the Protein Production Facility.

Special recognition goes to Daniel-Adriano Silva for his essential computational protocols and his experimental assistance, but also for his seemingly limitless enthusiasm and curiosity.

I am also deeply indebted to K. Christopher Garcia, Jamie Spangler, and other members of the Garcia Lab, as well as to Marion Pepper and her lab for their collaboration.

Finally, I would like to thank my family and friends for their support; my supervisory committee of Dr. James Carothers, Dr. Wendy Thomas, Dr. Valerie Daggett, Dr. Kim Woodrow, Dr. Marion Pepper, and Dr. Julie Overbaugh for their guidance; and most of all, to Dr. David

Baker for providing me with his mentorship and the opportunity to work in such an amazing lab over the years!

Table of Contents

Introduction and Background

Protein-protein interactions are central to many processes in biology, including host-pathogen recognition and cell signaling. The rapidly increasing number of deposited protein complex structures has allowed structural biologists to better understand the relationship between protein structure and function. Meanwhile, rational design efforts have matured over the course of several decades; with the aid of computational tools, it has become possible to rationally design protein binders to a specific interface in order to inhibit or otherwise modify the native interaction.

Computational Protein Design

The field of rational protein design has undergone tremendous progress over the course of several decades. Rational design efforts from first principles without the use of sophisticated computational methods saw the successful design of *de novo* helices, sheets, and even bundles, onto which catalytic activity could be introduced.[1–3] In addition, the use of programs such as DEZYMER[4–6] and ORBIT[7] has guided the design of protein motifs from novel sequences as well as biosensors and enzymes. In 2007, Liu *et al.* used a combination of programs, including SCAP[8] and CHARMM[9] to guide the grafting of key interacting residues to achieve a high-affinity protein-protein interface interaction.[10]

Rosetta

The ROSETTA molecular modeling software consists of a number of applications which can be used in the structural prediction and design of proteins as well as other macromolecules.[11,12] Successful designs efforts using ROSETTA encompass novel folds[13], geometrically idealized versions of existing folds[14], helical bundles[15], repeat proteins[16,17], and

symmetric assemblies[18,19]. ROSETTA proteins have been designed for binding to DNA[20,21], binding to various ligands[22–24], and even catalysis of a diverse range of substrates[25–29].

ROSETTA has also been successfully used in protein interface design targets, including influenza virus hemagglutinin[30,31], immunoglobulin G[32], and Epstein-Barr viral Bcl-2 protein[33]. Moreover, these protein interface design efforts employed a variety of methods including "hotspot-centric" centric *de novo* design[34], in which the binding interactions were explicitly designed into optimized "hotspot" residues; a side-chain grafting approach[35], in which motifs of interest were placed onto naturally existing proteins; and a new Fold From Loops procedure[36] in which new protein structures were built *de novo* using only a functional input motif and a target topology as guide. The work covered in this thesis in designing in an interleukin-2 mimetic makes use of existing methods and principles used in previous protein interface design efforts, but also features a new computational method in which the entire protein is designed *de novo* to give a geometrically idealized variant with regular secondary structural elements and shorter loops.

Interleukin-2

IL-2 is an immunoregulatory cytokine that promotes the proliferation, differentiation, and survival of mature T and B cells. It is produced primarily by $CD4^+$ T cells following antigen stimulation[37], but to a lesser extent by $CD8^+$ cells, NK T cells, activated dendritic cells, and mast cells as well[38].

IL-2 binds to cell surface receptors formed by various combinations of three IL-2R subunits: IL-2Rα (CD25), IL-2Rβ (CD122), and IL-2Rγ (CD125 or the common cytokine receptor γ chain, γ_c), as illustrated in Figure I-1. The interaction between formed between IL-2 and IL-2Rα is termed the "low-affinity" complex. The combination of IL-2Rβ and γ_c forms the

"intermediate-affinity" receptor, although IL-2 binds weakly to IL-2Rβ alone and does not bind at all to γ_c alone; the intermediate affinity complex is the minimal functional form of the receptor, as binding event causes IL-2Rβ and γ_c to heterodimerize, unleashing intracellular signaling pathways via the JAK1 and JAK3 kinases associated with the intracellular domains of IL-2Rβ and γ_c, respectively. The combination of IL-2Rα, IL-2Rβ, and γ_c forms the "high-affinity" complex, with IL-2Rα causing a conformational change in IL-2 that allows the cytokine to better engage IL-2Rβ and γ_c. IL-2 exerts pleiotropic effects through binding to different receptor complexes, with components differentially expressed across different cell types. For instance, because regulator T cells and activated T effector cells express high levels of IL-2Rα in addition to IL-2Rβ and γ_c, they are preferentially stimulated over NK and memory $CD8^+$ T cells which generally express low levels of IL-2Rα.

IL-2 has been explored for a variety of potential applications, including as an antitumor immunotherapy[37] and as an adjuvant in anti-retroviral therapies and other treatments for immunocompromised patients[39]. Due to its ability to stimulate regulatory T cells, IL-2 has also been used for suppression of autoimmune diseases such as Type 1 diabetes[40] and myasthenia gravis[41] as well as infectious diseases such as tuberculosis[42], where the ability to modulate immune response prevents destructive immunopathology without excessively suppressing protective immunity. Finally, it has been proposed that IL-2 inhibitors could serve useful immunosuppressive purposes[38] to complement current inhibitors that currently target the downstream signaling pathways mediated through IL-2 activation of JAK1[43] and JAK3[44].

Using directed evolution approaches, Levin *et al.* have engineered IL-2 "superkines" which exhibit dramatically increased binding affinity for the IL-2Rβ/γ_c intermediate-affinity receptor in the absence of IL-2Rα (Figure I-2).[45] They demonstrate that these IL-2 superkines

exhibit potent antitumor efficacy in four different tumor models, with far less pulmonary edema than induced by wild-type IL-2, which is attributed to preferential stimulation of cytotoxic T cells over regulatory T cells made possible by the reduced dependence on IL-2Rα. Furthermore, subsequent IL-2 receptor "signaling clamps" based on these superkines have weakened binding to γ_c, thereby serving as partial agonists or non-agonists that are immunomodulatory or immunosuppressive.[46] In short, conventional protein engineering methods have been employed to develop a wide range of cytokines with different, desirable applications depending on clinical context.

Computationally designed interleukin-2 mimetics

In the course of my doctoral work, I have designed and optimized a wide range of IL-2 mimetics which fall into two broad categories. The "three-helix designs" are derived from a naturally occurring protein of with some structural similarity but unrelated function to that of the IL-2, with mutations made to the existing protein to repurposes it for binding to the IL-2 receptors, using methods based on existing side-chain grafting approaches previously employed by Azoitei *et al*.[47] The "four-helix designs" are idealized, *de novo* creations that recapture the structural features of IL-2 that allow it to bind the intermediate affinity IL-2Rβ/γ_c complex, designed using new computational methods. Neither the three-helix nor the four-helix designs maintain any structural analog to that of helix B in IL-2, which forms the contact with IL-2Rα, and as such, which allows these mimetics to bind the intermediate affinity complex with even less dependence on IL-2Rα. These mimetics are developed in hopes of forming the basis of new therapeutics, and a wide range of variants belonging of such mimetics are engineered in hopes of better elucidating the functional effects of binding affinity on the IL-2 pathway.

Figures

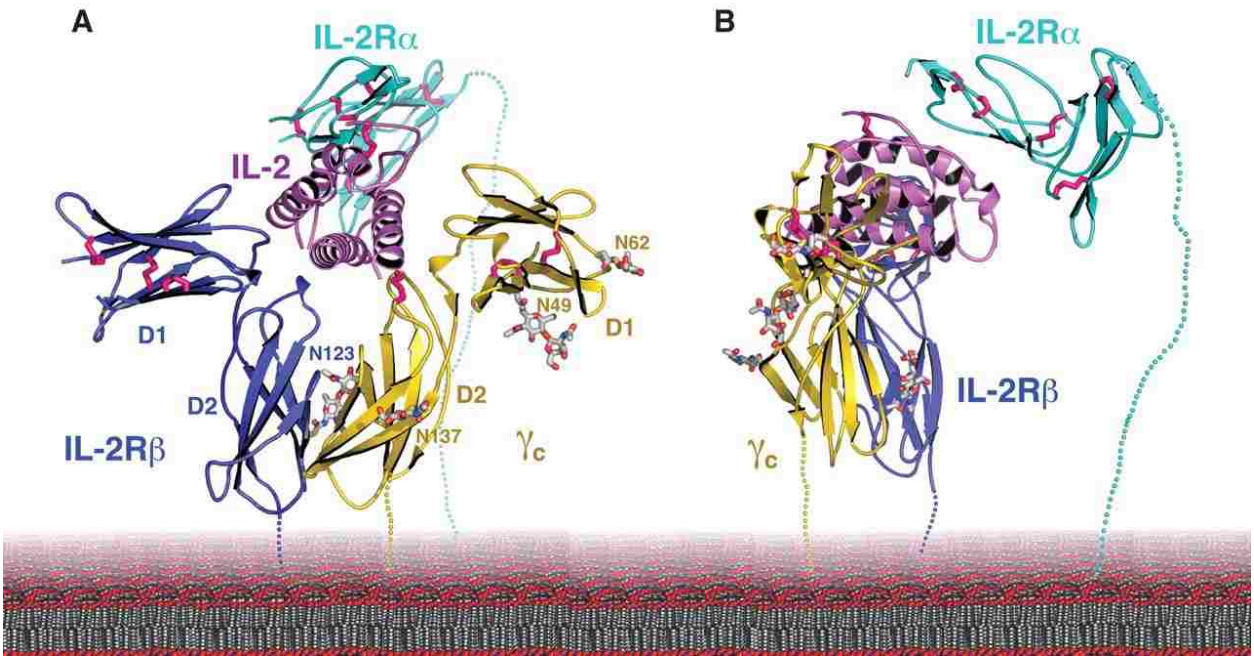


Figure I-1: IL-2 quaternary complex. (a) The high-affinity receptor complex consists of the interaction between IL-2 (violet), IL-2Rα (cyan), IL-2Rβ (blue), and γ_c (gold); (b) side view of the same. Image from Wang *et al*.[48]

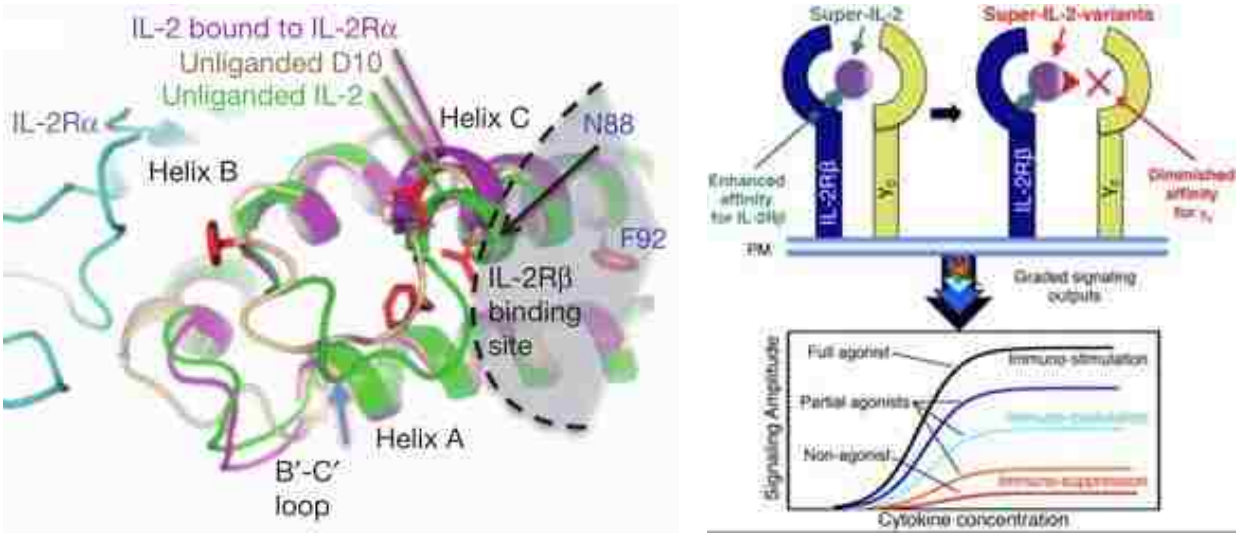


Figure I-2: IL-2 superkines or receptor signaling clamps allow for more precise tuning than are possible with wild-type IL-2 cytokine. *Left*—IL-2 superkine D10 contains core mutations that stabilize a conformation similar to that of IL-2 in its liganded conformation, thereby allowing it to better bind IL-2Rβ even in the absence of IL-2Rα.[45] *Right*—receptor signaling clamps with mutations that diminish the affinity for γ_c allow for precise control of signaling output.[46]

Computational Design of Interleukin-2 Mimetics

Background

Interleukin-2 (IL-2) is an immunoregulatory cytokine that promotes the proliferation, differentiation, and survival of mature T and B cells. As a cytokine with paradoxical effects on the immune response, promoting both immunostimulatory effector cells and immunosuppressive regulatory cells, IL-2 is a target of significant therapeutic importance.[37–44] The molecular basis of these diverse effects is explained by its interaction with its receptors: through the formation of an intermediate-affinity ternary complex consisting of IL-2, IL-2Rβ, and γ_c, the cytokine can initiate signaling pathways via intracellular JAK kinases complexed with the IL-2Rβ and γ_c receptor subunits. Alternatively, the signal can also be initiated through the formation of a high-affinity quaternary complex that also includes IL-2Rα, a non-signaling subunit that is more highly expressed on regulatory T cells.

Recently developed IL-2 "superkine" mutants have enhanced binding affinity for IL-2Rβ/γ_c even in the absence of IL-2Rα, inducing superior expansion of cytotoxic T cells, leading to improved antitumor responses *in vivo*, with less expansion of T regulatory cells and reduced pulmonary edema.[45] Moreover, variants of such superkines have been engineered for weaker interactions with γ_c, allowing them to be tuned as partial agonists for immunomodulation or even non-agonists for immunosuppression.[46] The highest affinity superkines characterized differ from wild-type IL-2 by only 5-6 amino acids and are nearly structurally identical. Computationally designed proteins that preserve only the structural elements necessary for binding IL-2Rβ/γ_c can provide not only a validation of our computational design methods, but also a cytokine mimetic that is truly independent of any interactions with IL-2Rα.

Results

Computational Design by MotifGraft

Two computational design methods were employed in the development of first-generation binders against IL-2β/γ_c. The first method (MotifGraft) featured a side-chain grafting approach for transplanting binding motifs onto naturally occurring scaffolds, previously employed in the presentation of HIV epitopes[35,47] and in the design of inhibitors of Epstein-Barr Viral Bcl-2 protein[33]. Three helices were used in various combinations to serve as input for MotifGraft. During the construction of input motifs, helix A was defined as encompassing residues from Lys-8 to Gly-27 of PDB 2B5I, chain A, forming contacts against both IL-2Rβ and γ_c; helix C encompassing residues from Arg-83 to Lys-97, contacting IL-2Rβ; input helix D encompassing residues from Val-115 to Thr-131, contacting γ_c. Three separate input motifs were defined as containing helices A and D; helices A and C; or helices A, C, and D. These input motifs were run against a small pre-existing scaffold set of 1046 entries from the PDB (Table S1-1), as well as against a list of 23 scaffolds (Table S1-2) which were identified as compatible with one of the three input motifs using a TM-align[49] search against a more expansive set of scaffolds. MotifGraft yielded a total of 262 designs, of which 19 matched the motif for helices A and C and 243 matched helices A and D. No MotifGraft designs matched motifs for all three helices. After extensive consideration involving both automated filtering and manual inspection, thirty sequences (corresponding to 15 unique designs and variants thereof) were selected for experimental testing (Table 1-1). Of 15 these unique designs, only one (corresponding to the four sequences Rbg01, Rbg02, Rbg03, and Rbg29) matched helices A and C; the remainder matched helices A and D (Figure 1-1). Of these 30 binders, only 24 were tested, but none exhibited binding on yeast display.

Focused Library Selection and Emergence of Rbg32

During the semi-automated process of introducing beneficial mutations to the original 30 designs, several candidate mutations were identified as being potentially beneficial to the design in terms of different metrics (such as binding energy, total pose energy, buried solvent-accessible surface area, etc.). Seven different unique designs (Rbg03, Rbg12, Rbg16, Rbg17, Rbg18, Rbg21, and Rbg25) were selected for diversification based on a combination of these mutations, resulting in seven "focused" libraries each containing 7 to 12 mutated residues representing an amino acid diversity of 1024 to 65536.

Binding signal emerged from the selection of the library based on Rbg21 (Figure 1-2), a design with input motif containing helices A and D. Sanger sequencing results revealed strong convergence towards a common sequence, subsequently labeled as Rbg32. Notably, many of the mutations in which Rbg21 differed from Rbg32 are found on helix A in which the latter contains residues that are found on wild-type human IL-2 (Figure 1-3) , suggesting that more conservative design of Rbg21 that recapitulates the structural features of IL-2 may have yielded more successful initial designs. However, residues on positions that would correspond to helix D were not as conserved in this regard.

The scaffold protein for Rbg32 is the mitochondrial interacting domain of an AAA ATPase Vps4 from *Sulfolobus solfataricus* (PDB: 2v6y, chain A) Structurally, Vps4 protein consists of three "open" helices, interacting with Vps2 to complete a four-helix bundle.[50] However, in the computational model of the interaction between Rbg32/IL-2Rβ/γ_c, the "missing" fourth helix that would be represented by Vps2 is instead distal to junction formed by IL-2Rβ/γ_c. While the absence of such a helix suggests that the binding interaction between Rbg32 and IL-2Rβ/γ_c is likely to be unaffected by the presence of IL-2Rα, it also suggests that the resulting

Rbg32 designed binder may not be especially stable, as the scaffold protein has been removed from its native biochemical context in which Vps2 was present. As such, single disulfide bonds were engineered into six Rbg32 variants (Table 1-2) to confer better stability. Two such variants (Rbg32.4 and Rbg32.5) exhibited improved binding signal, presumably due to enhanced stability, and cysteine mutations in from both Rbg32.4 and Rbg32.5 were introduced into the Rbg32.7 variant containing two disulfide bonds (Figure 1-4). However, as can be seen in Figure 1-9, yeast surface titrations against biotinylated human IL-2Rβ/γ_c (b-hIL-2Rβγ) indicate that Rbg32.4, Rbg32.5, and Rbg32.7 bind very weakly (K_d ~ 1-10 μM) compared to IL-2 (K_d = 240-260 nM).

Computational Design of De Novo *Binders with Idealized Geometry*

A second computational approach employed in the design effort featured a new idealization protocol in which the entire protein sequence was designed computationally to recapitulate the three-dimensional arrangement of the secondary structures of IL-2. This yielded a total of 210 candidate designs, of which 62 were explicitly restricted from having methionine residues in the core and 148 were allowed to have methionine. After testing with *ab initio* forward folding tests, four designs (Rbg33, Rbg36, Rbg40, and Rbg42) were selected for experimental testing (Table 1-3). Structurally, these designs are 87-residue, four-helix bundles that closely recapitulate the topology of wild-type human IL-2, especially at helices A, C, and D, where IL-2 binds IL-2Rβ/γ_c (Figure 1-5). These designed mimetics do not contain the loops that connect helices A to B or helices C to D on IL-2, which run antiparallel and parallel to helix B, respectively. As a result, the mimetics less faithfully recapitulates the structure of helix B. Because helix B and these loops form the interface that would bind IL-2Rα, these *de novo* designs are likely to bind IL-2Rβ/γ_c with minimal dependence on IL-2Rα (Figure 1-6). In

addition to these four designs, eight additional designs representing variants of the above with disulfide bonds engineered were also ordered for experimental testing.

Experimental Characterization of Four-Helix Binders

Of the twelve four-helix designs ordered, ten (all but Rbg33 and Rbg34) exhibited binding signal at 1 μM b-IL-2Rβγ on yeast display. Six exhibited binding signal at 1 nM b-IL-2Rβγ (Figure 1-7), of which Rbg40 only did not contain a disulfide bond. Rbg40, Rbg41, and Rbg43 were chosen for expression and characterization by circular dichroism (Figure 1-8). Both Rbg40 and Rbg41 exhibit some loss of native conformation when heated to 95°C but retain most of that character when cooled back down to 25°C, whereas Rbg43 (not shown) is marked by more significant loss of structural character. Rbg40 and Rbg41 also show cooperative unfolding under chemical denaturation, with Rbg40 exhibiting greatest stability. Neither Rbg40 nor Rbg41 are completely unfolded at a maximum temperature of 95°C in a thermal melt.

Yeast surface titrations against b-hIL-2Rβγ show that while Rbg32.4, Rbg32.5, and Rbg32.7 have little or no binding affinity, Rbg40, Rbg41, and Rbg43 have binding affinities comparable to that of IL-2 (Figure 1-9), with enhanced cooperativity for the ternary complex compensating for slightly weaker binding to the hIL-2Rβ receptor subunit (Figure 1-10). Moreover, in addition to binding b-hIL-2Rβγ, the four-helix binders have been demonstrated by Spangler[51] to be capable of stimulating STAT5 phosphorylation in YT-1 natural killer cells with half maximal effective concentrations comparable to those of IL-2 (Figure 1-11), demonstrating that designed mimetics are capable of inducing biological signaling.

Methods

Preparation of computational models

The structure of human IL-2, IL-2Rβ, and γ_c in quaternary complex (along with IL-2Rα) was taken from the Protein Data Bank (PDB: 2B5I).[48] The protein structures were energetically minimized in full-atom mode, subject to coordinate constraints under a harmonic function ("relax with coordinate constraints", Protocol S1-1). In addition, a separate PDB ("IL-2 file") consisting of just chain A from 2B5I (IL-2) was also generated using relax with coordinate constraints. A third PDB file ("target file") was generated consisting of chains B and C from 2B5I (IL-2Rβ and γ_c, respectively) renumbered and renamed as a single chain.

Computational design of interleukin-2 mimetics

Helices A and C; helices A and D; or helices A, C, and D from the IL-2 file above were used as input motifs. A pre-existing scaffold set (Table S1-1) consisting of 1046 entries was supplemented with 23 entries identified by TM-align[49] as potential structurally compatible scaffolds (Table S1-2). Candidate scaffolds were run through a modified version of Multigraft Match[35], originally described by Correia *et al.*, adapted for RosettaScripts[52], using the target file described above (Protocol S1-2). Briefly, compatible scaffolds were matched for surface-exposed segments with conformation similar to those of the input motifs. Sidechains from input motif were transferred onto the candidate scaffold, then filtered for backbone clashes with the target and for binding energy. Amino acids within the scaffold but outside the input motif were designed against the target.

Four-helix IL-2 mimetics based on the original IL-2 file were idealized using a new computational method (Protocol S1-3).[53] Briefly, overlapping 4-mer fragments from the input

were clustered by torsion angle from the vall database of all known fragments[54] from the PDB. Transition count matrices of clusters were used to idealize secondary structure. Loops were introduced to connect secondary structures using a novel loop building method based on assembly of 4-mers. Designs were then subject to sequence optimization followed by forward folding simulations.

Addition of disulfide bonds

Single disulfide bonds were added to Rbg32, Rbg33, Rbg36, Rbg40, and Rbg42 using the disulfide-building option of RemodelMover in RosettaScripts (Protocol S1-4).[52] Briefly, six-dimensional rotation-translation vectors between residue backbones for all combinations of amino acid pairs in the input sequence were compared against a database of such vectors for all known disulfides.

Two disulfide bonds were added to Rbg32 to yield Rbg32.7 by combining the cysteine mutations from Rbg32.4 and Rbg32.5.

Yeast cell transformation

Designs Rbg01 to Rbg30 were reverse translated into genes optimized for *Escherichia coli* using DNAWorks[55], ordered as synthetic DNA from Gen9; designs Rbg32, Rbg32.1 to Rbg32.8, and Rbg33 to Rbg44 were ordered as gBlock gene fragments from Integrated DNA Technologies. Genes were transformed by heat shock into *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector (a modified version of the pCTCON2 yeast display vector originally described by Chao *et al*.), double digested by NdeI and XhoI restriction enzymes.[56] Sanger sequencing of colony PCR products were used to verify successful homologous recombination.

Focused libraries corresponding to candidate MotifGraft-based designs Rbg03, Rbg12, Rbg17, Rbg18, Rbg21, Rbg25, and Rbg30 were created by assembly PCR using oligonucleotides shown in Table S1-3. Genes were transformed by electroporation into conditioned *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector, using the protocol previously described by Benatuil *et al*.[57]

Yeast display screening

Yeast cells were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm, then induced in 1 mL SGCAA for expression at 30°C for 16 h. Screens were conducted using either tetramerized or non-tetramerized human IL-2Rβ and γ_c receptors, expressed as acid/base zipper heterodimers with C-terminal BAP tag (Table S1-4) to facilitate biotinylation (b-hIL-2Rβγ), provided as a gift from the Garcia Lab. For initial screening assays using tetramerized b-hIL-2Rβγ, tetramers were formed by incubating 1 μM b-hIL-2Rβγ with 250 nM streptavidin, R-phycoerythrin conjugate (SAPE, Immunology Consultants Laboratory) in phosphate buffered saline supplemented with 0.5% BSA and 2 mM EDTA (PBE) for 15 min on ice. Yeast cells were then incubated with tetramerized b-hIL-2Rβγ for 2 h at room temperature in PBE, then labeled with 1 μg FITC conjugated chicken anti-C-Myc (anti-C-Myc-FITC, Life Technologies) in 50 μL PBE per 1×10^6 cells for 15 min on ice. For subsequent screening assays using non-tetramerized b-hIL-2Rβγ, cells were incubated with b-hIL-2Rβγ for 2 h at room temperature in PBE, then labeled with 1 μg SAPE and 1 μg anti-C-Myc-FITC in 50 μL PBE per 1×10^6 cells for 15 min on ice. Measurements were performed on an Accuri C6 flow cytometer.

Fluorescence activated-cell sorting

Yeast cells containing focused libraries were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm, induced in 1 mL SGCAA for expression at 30°C

for 16 h, incubated with 1 μM b-hIL-2Rβγ pre-tetramerized with 250 nM SAPE for 2 h at room temperature in PBE, labeled with 1 μg anti-C-Myc-FITC in 50 μL PBE per 1×10^6 cells for 15 min on ice, and collected in 1 mL c-Trp-Ura media based on a combination of gates based on FSC, SSC, PE, and FITC signals. Selection process was repeated for a total of four rounds. Colony PCR was performed on cells collected from final round of selection, and Sanger sequencing was performed on colony PCR products to determine the sequence identity of a sample of 12 clones from the final enriched library. FACS experiments were performed on the BD Influx cell sorter from BD Biosciences.

Recombinant expression

Plasmids containing genes encoding designs of interest in pETCON yeast expression vector were extracted from yeast EBY100 cells using Zymolyase from Zymo Research. Genes of interest were amplified from pETCON with primers shown in Table S1-5, inserted into pET29b *E. coli* expression vector using Gibson assembly[58], and transformed into *E. coli* XL10-gold strain cells by heat shock. Plasmids containing genes encoding designs of interest in pET29b vector were extracted from *E. coli* XL10-gold cells by Miniprep (Qiagen) and transformed into *E. coli* BL21(DE3) strain cells. Cells were grown at 37°C in 0.5 L Terrific Broth to an optical density of 0.6 at 600 nm and induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for expression at 18°C for 6 h. Cell pellets were harvested by centrifugation and lysed by sonication. Protein was purified by Ni-NTA chromatography (eluted in PBS with 250 mM imidazole), verified by SDS-PAGE and mass spectrometry, and again purified for monomers via size exclusion chromatography (eluted in PBS).

Circular dichroism

Circular dichroism wavelength scan measurements were conducted on an Aviv Circular Dichroism Spectrometer, Model 420. Wavelength scan spectra were recorded from 260 nm to 195 nm in increments of 1 nm at 25°C and 95°C, and then repeated at 25°C. Proteins were diluted to 0.1 mg/mL in PBS, pH 7.4 buffer, for normal conditions; in PBS with 1 mM TCEP, pH 7.4 buffer, for reducing conditions; or in PBS with 1 mM potassium ferricyanide, pH 7.4 buffer, for oxidizing conditions. Measurements were conducted in a cuvette with pathlength 1 mm.

Thermal denaturation curves were conducted in 2°C steps with a heating rate of 2°C/min, with the absorption signal measured at 222 nm. Protein samples were diluted to 0.1 mg/mL in PBS, pH 7.4 buffer, for normal conditions; in PBS with 1 mM TCEP, pH 7.4 buffer, for reducing conditions; or in PBS with 1 mM potassium ferricyanide, pH 7.4 buffer, for oxidizing conditions. Measurements were conducted in a cuvette with pathlength 1 mm.

Chemical denaturation experiments were conducted using titration by guanidinium hydrochloride (GuHCl). Protein samples without disulfide bonds were prepared in 0.1 mg/mL in PBS, pH 7.4 buffer. Protein samples with disulfide bonds were prepared in 0.1 mg/mL in both PBS, pH 7.4 buffer, and in PBS with 1 mM TCEP, pH 7.4 buffer. Samples were titrated with GuHCl ranging from 0 M to 4 M, in increments of 0.5 M. Measurements were conducted in a cuvette with pathlength 1 cm using a Hamilton Microlab 500 Series Diluter and Dispenser.

Discussion

The disparity in success rate of the putative binders designed by the MotifGraft approach and those designed by the idealization protocol highlights the advantages of the latter. Naturally occurring proteins offer a large but limited repertoire of structures from which to design.

Moreover, most mutations, including those intended to confer functional properties such as improved binding, are destabilizing, and naturally existing proteins are generally selected only for the minimum stability needed for biological function.[59] In contrast, *de novo* approaches provides the limitless canvas of any biophysically accessible protein structure, limited only by our computational abilities, and furthermore, *de novo* proteins can be designed with greater stability than anything found in nature.[13–15]

Nonetheless, the Rbg32 binder that resulted from the MotifGraft approach has distinctive advantages. γ_c is a cytokine receptor subunit that is common to a large family of receptor complexes, but IL-2 binds γ_c only very weakly.[48] Because the input helix D used in the MotifGraft approach interacts at the Rbg32/γ_c interface, an optimized version of Rbg32 may be a γ_c-specific binder or the basis of such a binder when used as a template in conjunction with methods such as the idealization protocol. Indeed, conventional sequencing results from the selection of the lib_Rbg21 focused library (Figure 1-2) indicate that most of the mutations at helix A (lying at the junction between IL-2Rβ/γ_c) were reversions to IL-2, but that mutations contacting γ_c tended to adopt different identities, suggesting that such residues are less conserved. Indeed, because γ_c is a common receptor to many cytokines responsible for so many biological functions when heterodimerized with other receptor subunits, IL-2 and other cytokines may not have been selected during the course of natural evolution for extremely strong binding to γ_c.

The idealized four-helix *de novo* designs feature binding affinities comparable to that of IL-2 cytokine and have been capable of stimulating STAT5 phosphorylation despite no experimental optimization. Rbg40 will serve as the starting point for future directed evolution experiments for improved binding, as circular dichroism measurements suggest that it is the most

stable of the characterized variants, and the absence of any disulfide bonds facilitates recombinant expression under a wider range of conditions. Recombinant IL-2 mimetics based on Rbg40 may have several advantages over existing cytokines, including smaller size, greater ease of expression, and improved thermostability. Moreover, excessive loops in IL-2 that are vestigial remnants of evolution have been excised during the computational design process, further stabilizing the mimetics against unwanted interactions such as proteases. Finally, helix B and the two extensive loops that connect helices A and B and helices C and D on IL-2 are replaced with a single helix on Rbg40, thus eliminating the interface for any potential interaction with IL-2Rα. As a result, IL-2 mimetics based on Rbg40 are likely even less dependent on IL-2Rα than even recently evolved superkines[45] for binding to IL-2Rβ/γ_c.

Tables and Figures

Design	**Scaffold**	**Sequence Identity**
Rbg01	1TQGA	EYLGVFADETKEYLQNLNDTLLELEKNPEDMELINEAFRAALSLLGMAGTMGFSSLLK VCIALENAADKARNSEIKTTSDALDAARAGVEFITRMVDKIVS
Rbg02	1TQGA	EYLGVFADETKELLQNLNDTLLELEKNPEDMELINEAFRAALSLLGMAGTMGFSSLLL VAIALLLAADKARNSEIKTTSDALDAARAIVEFITRFVDKIVS
Rbg03	1TQGA	EYLGVFADETKELLQNLNDTLLELEKNPEDMELINEAFRAALTLLGMAGTMGFSSLLK ACIALLNAADAARNSEIKTTSDALDAARAIVEFITRFVDKIVS
Rbg04	4A5XA	PQSTAAATALKRAVELDSESRYPQAAACAQEGINLLAQVIRGTKDNTKRCNLREKASK AMDRLEDIQKYLDQE
Rbg05*	3EABA	GSMEAERARVWHKQAFEWASIALRIDEDEKAGQKEQAAEWLKKGAEAAQAGIAVIITG QGEQCERARRLQAILMANLVDAQERLQLIE
Rbg06*	3EABA	GSMEAERVRVFHKQAFEAISIALRIDEDEKAGQKEQAVEAAKVGLEALRQGIAVIVTG QGEQCERARRLQAKAMTNAVMAQARIALLE
Rbg07	2W2UA	SAQVMLEEMARWAAIAAVKADKEGNAEEAARAAKAALEALRQLASLYRDGSTAAIAEQ MANEAQRRIAVLKELI
Rbg08	2V6XA	DWLTKGIELAQKAIDLDTATQYEEAYTAAQNGIDAAQAAIAYAKNPRSKEIARAKLTD LLTRAEQIKKHLESEEAN
Rbg09*	2V6XA	DFLTKGIEWLQKAIDLDTATQYEEAATAYQAGISYAALALKYEKNPKSKDLIRAKITE AIARLEDIQKHLESEEAN
Rbg10	2V6XA	DFLTKGIEWLQKAIDLDTATQYEEAATAYQAGISYAALALKYEKNPKSKDLIRAKIRE AIARLEDIQKHLESEEAN
Rbg11	2V6XA	DFLTKGIEWLQKAIDLDTATQYEEAATAYQAGISYAALALKYEKNPKSKDLIRAKILE AIARLMDIQKHLESEEAN
Rbg12	2QSBA	VRVDQNLFNEVMQLLDELSQDITSPKNVRKLAQDAAAKLSQENESLDLACATAISMAQ EAIADPNVPARGRTDLIAILSALEAIS
Rbg13	2QSBA	VRVDQNLFNEVMNLLDELSQDITSPKNVRKLAQDAAAKLSQENESLDLACATAISMAQ EAIADPNVPARGRRDLIAILSALEAIS

Rbg14*	2V6XA	DWLTKGIELAQKAIDLDTATQYEEAYTAAQNGIDAAQAAIAYAKNPRSKDIARAKLTD LLTRAEQIKKHLESEEAN
Rbg15*	2V6YB	SAQVMLEDLARAAAIAAVKADKEGKVCEAVRAYKAALEAARQLIVLYPESVARTAYEQ MANEAQARISYLEKVL
Rbg16	1X91A	SSEMSTICDKTLNPSFCLKFLNTKFASANLQALAKTTLDSTQARATQAIAKAQAAIAG GVDPESKLAYRACLDALQNAIGNAEEAFEHAASGDGMGANMKVSAALDGADWCLDALS RLRSVDSSAVNNAKTLKNLCGIALVIANMLPRN
Rbg17	1Z6OM	HRSCRNSMRQQIQMAIGASLQALAMGAHASKDVVNRPGVAQLAFDAASEEREHAMKLI ELLLMRGELTNDVSSLLQVRPPTRTSWKGGVEALEHAASMEQAITESARNVIKACEDD SEFNAYHLADYLTGDIVEEQIAGLRDVQGKASTLKKLMDRHEALGEFIFAKKLLGIDV
Rbg18	3EABA	GSMEAERVRVWHKQAFEWISIALRIDEDAKAGQKEQAIEWYKKGIEALQAAIAVIVTG QGEQCERARRMQKILMANLVDALDRLQLLE
Rbg19	2W2UA	SAQVMLEEMARKYAINAVKADKEGNAEEAITNAKKAAEAAQWLIALYKDGSTARIYEA MFNDLKRRIEVLKELI
Rbg20	2W2UA	SAQVMLEEMARKYAINAVKADKEGNAEEAITNAKKAAEAAQWLIALYWDGSTAAILIA MLSDLQWRIEVLKELI
Rbg21	2V6YA	SAQVMLEDLARELAIAAVKADKEGKVEDAATYYKAALEAVRQIIVLYPESVARTAYEQ MANEAQKRIAYLEKVL
Rbg22	3EABA	GSMEAERVRVWHKQAFEWISIALRIDEDAKAGQKEQAIEWYKKGIAALQAAIAVIVTG QGEQCERARRMQAILMANLVDALDRLQLLE
Rbg23	1Z6OM	HRSCRNSMRQQIQMAIGASLQALAMGAHASKDVVNRPGVAQLAFDAASEEREHAMKLI ELLLMRGELTNDVSSLLQVRPPTRTSWKGGVEALEHAASMEQAIIASARNVIKACEDD SEFNAYHLADYLTGDILREQIAGLRDVQGKASTLKKLMDRHEALGEFIFAKKLLGIDV
Rbg24	2V6YA	SAQVMLEDLAREAAIAAVKADKEGKVEDAARYYKAALEAVRQIIVLYPESVARTAYEQ MANEAQKRIAYLQKVL
Rbg25	3EABA	GSMEAERVRVWHKQAFEAISIALRIDEDEKAGQKEQAAEAAKAGLEAARQGIAVIVTG QGEQCERARRLQAKMMTNVVMAQARISLLE
Rbg26	3EABA	GSMEAERVRVFHKQAFEAISIALRIDEDEKAGQKEQAVEAAKAGLEALRQGIAVIVTG QGEQCERARRLQAKAMTNAVKAQARIALLE
Rbg27	3EABA	GSMEAERARVWHKQAFEWASIALRIDEDEKAGQKEQAAEWLKKGAEAAQAAIAVIITG QGEQCERARRLQAILMANLVDAQERLQLIE
Rbg28	4A5XA	PQSTAAATALKRAVELDSESRYPQAAACAQEGINLLAQVIRGTKDNTKRCNLREKASK AMDRLEDIRKYLDQE
Rbg29	1TQGA	EYLGVFADETKEYLQNLNDTLLELEKNPEDMELINEAFRAALSLLGMAGTMGFSSLLK ACIALENAADKARNSEIKTTSDALDAARAGVEFITRMVDKIVS
Rbg30*	1X91A	SSEMSTICDKTLNPSFCLKFLNTKFASANLQALAKTTLDSTQARATQAIAKAQAAIAG GVDPESKEAYRACLDALQSAIGNAEEAFEHAASGDGMGANMKVSAALDGADWCLDALS RLRSVDSSAVNNAKTLKNLCGIALVIANMLPRN

Table 1-1: Preliminary designs Rbg01 to Rbg30, corresponding scaffold PDB ID and chain letter used in Multigraft application, and amino acid sequence identity of design. Designs indicated in red served as the template for focused libraries, with corresponding red amino acid residues indicating positions replaced with degenerate codons. Asterisk (*) denotes designs not experimentally tested.

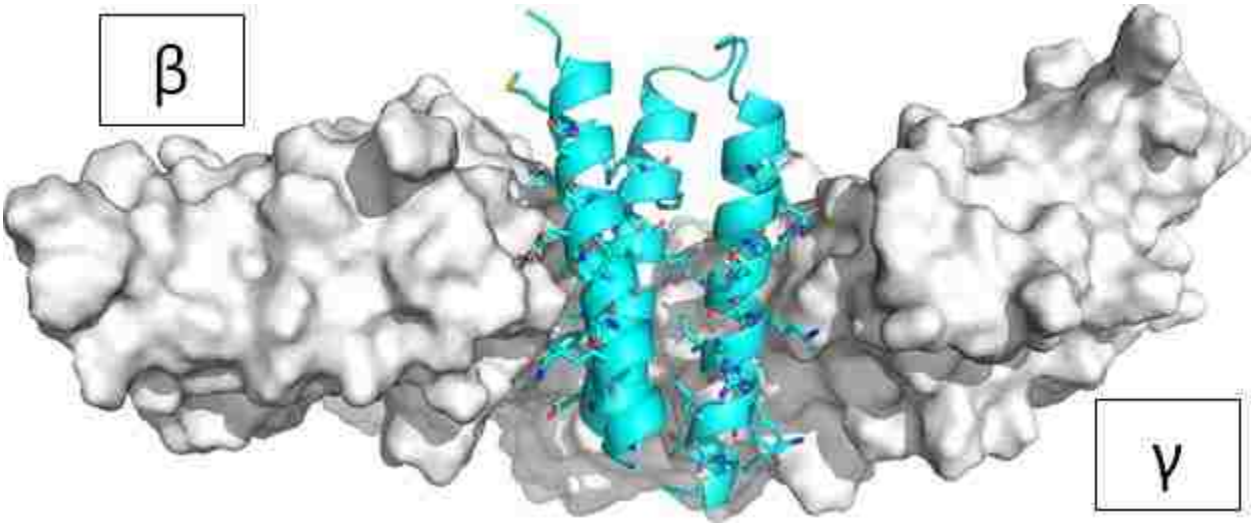
Design	**Sequence identity**
Rbg32	SAQVMLEDLARELAIAAVKADKEGKVEDAATYYEHALLDLQQIIVLYPESVARTAYEQMITEAQRRI ANLEKVL
Rbg32.1	SAQVMLEDLCRELAIAAVKADKEGKVEDAATYYEHALLCLQQIIVLYPESVARTAYEQMITEAQRRI ANLEKVL
Rbg32.2	SAQVMLEDLARECAIAAVKADKEGKVEDAATYYECALLDLQQIIVLYPESVARTAYEQMITEAQRRI ANLEKVL
Rbg32.3	SAQVMLEDLARECAIAAVKADKEGKVEDAATYYEHCLLDLQQIIVLYPESVARTAYEQMITEAQRRI ANLEKVL
Rbg32.4	SAQVMLEDLARELAIACVKADKEGKVEDAATYCEHALLDLQQIIVLYPESVARTAYEQMITEAQRRI ANLEKVL
Rbg32.5	SAQVMLEDLARELAIAAVKADKEGKVEDACTYYEHALLDLQQIIVLYPESVARTAYEQMITEAQRRI

	ANCEKVL
Rbg32.6	SAQVMLEDLARELAIAAVKADKEGKVEDAATYYEHACLDLQQIIVLYPESVARTAYEQMITECQRRI ANLEKVL
Rbg32.7	SAQVMLEDLARELAIACVKADKEGKVEDACTYCEHALLDLQQIIVLYPESVARTAYEQMITEAQRRI ANCEKVL

Table 1-2: Sequence identities of Rbg32, which emerged from selection of focused library lib_Rbg21, and its variants in which mutations to cysteine have been introduced to confer disulfide bonds.

Design	**Sequence Identity**
Rbg33	STKKWQLQAEHALLDWQMALNKSPEPNENLNRAITAAQSWISTGKIDLDKAEDIRRNSDQARREAEKRGI DVRDLISNAQVILLEAR
Rbg34	STKKWQLQAEHALLDWQMALNKSPEPNENLNRAITAAQSCISTGKCDLDKAEDIRRNSDQARREAEKRGI DVRDLISNAQVILLEAR
Rbg35	STKKWQLQAEHALLDWQMALNKSPEPNENLNRAITAAQSWISTGKIDCDKAEDIRRNSDQARREAEKRGI DVRDLISNAQVILLEAC
Rbg36	STKKLQLQAEHFLLDVQMILNESPEPNEELNRAITDAQSWISTGKIDLDRAEELARNLEKVRDEALKRGI DVRDLVSNAKVIALELK
Rbg37	STKKLQLQAEHFLLDVQMILNESPEPNEELNRCITDAQSWISTGKIDLDRAEECARNLEKVRDEALKRGI DVRDLVSNAKVIALELK
Rbg38	STKKLQLQAEHFLLDVQMILNESPEPNEELNRAITDAQSCISTGKCDLDRAEELARNLEKVRDEALKRGI DVRDLVSNAKVIALELK
Rbg39	STKKLQLQAEHFLLDVQMILNESPEPNEELNRAITDAQSWISTGKIDLDRAEELCRNLEKVRDEALKRGI DVRDLVSNACVIALELK
Rbg40	STKKLQLQAEHALLDAQMMLNRSPEPNEKLNRIITTMQSWISTGKIDLDGAKELAKEVEELRQEAEKRGI DVRDLASNLKVILLELA
Rbg41	STKKLQLQAEHALLDAQMMLNRSPEPNEKLNRIITTMQSCISTGKCDLDGAKELAKEVEELRQEAEKRGI DVRDLASNLKVILLELA
Rbg42	STKKIQLQLEHALLDVQMALNRSPEPNESLNRMITWLQSWISTGKIDLDNAQEMAKEAEKIRKEMEKRGI DVRDLISNIIVILLELS
Rbg43	STKKIQLQLEHALLDVQMALNRSPEPNESLNRMITWLQSCISTGKCDLDNAQEMAKEAEKIRKEMEKRGI DVRDLISNIIVILLELS
Rbg44	STKKIQLQLEHALLDVQMALNRSPEPNESLNRMITWLQSWISTGKIDLDNAQEMCKEAEKIRKEMEKRGI DVRDLISNICVILLELS

Table 1-3: Sequence identities of four-helix *de novo* designs Rbg33 to Rbg44. Rbg33, Rbg36, Rbg40, and Rbg42 are original *de novo* designs; the other eight are variants in which disulfide bonds have been engineered.



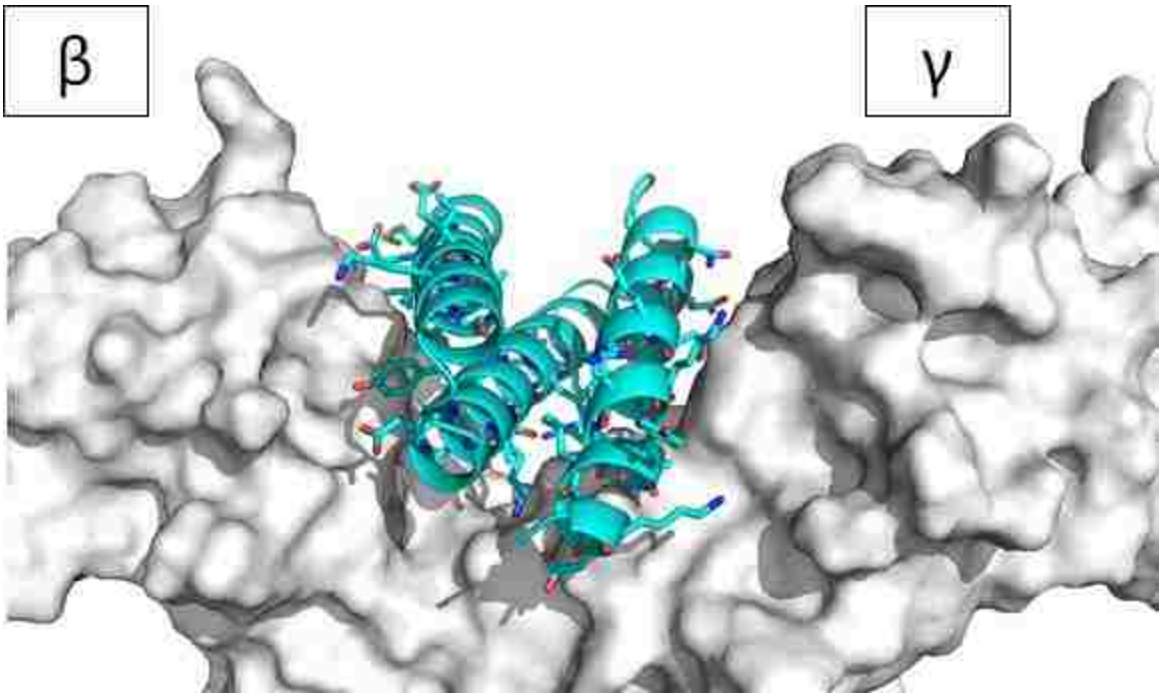


Figure 1-1: Computational model of Rbg21, a representative MotifGraft design. *Top*—Top view of Rbg21 (cyan), a three-helix design which lies at the interface of the IL-2Rβ/γ_c receptor. *Bottom*—Side view of the same. Computational models visualized using PyMOL.[60]

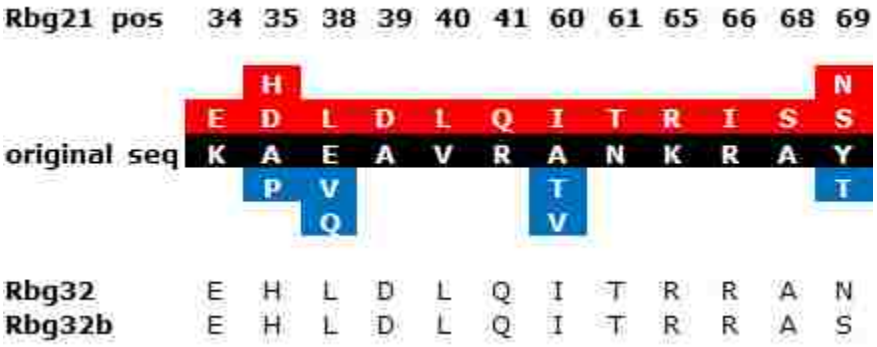


Figure 1-2: Focused library lib_Rbg21 contained mutations at 12 positions representing a diversity of 65536, converging on the sequence corresponding to Rbg32. Red background indicates desirable amino acids allowed by the designed codon; blue background indicates unintended amino acids allowed by the designed codon.

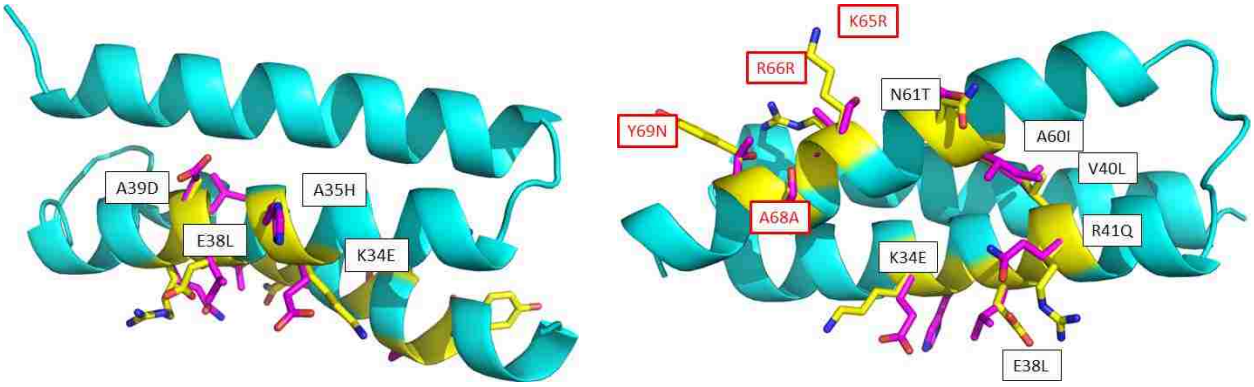


Figure 1-3: Comparison of Rbg21 (cyan) with wild-type human IL-2 (selected side chains shown in magenta sticks). Residues allowed to mutate in lib_Rbg21 are shown in yellow. *Left*—on helix A, residues on the original Rbg21 tend to revert to their corresponding identities on IL-2, as illustrated by K34E, A35H, E38L, and A39D observed on Rbg32. *Right*—mutation R41Q represents another reversion to IL-2 on helix A, but of the residues that correspond to helix D, only N61T represents a residue present on IL-2. K65R and Y69N, as well as the unmutated Arg-66 and Ala-68, are residues which differ from IL-2. Computational models visualized using PyMOL.[60]

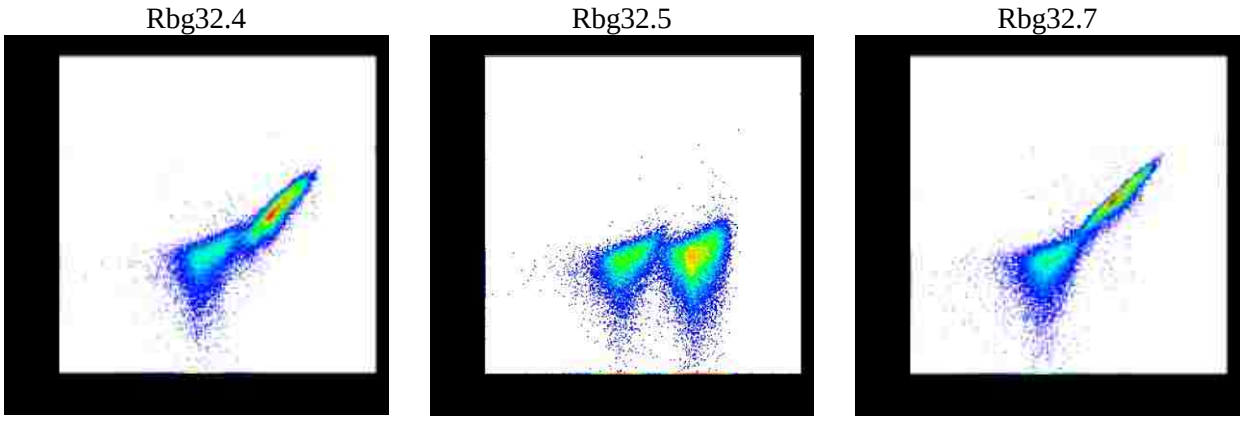


Figure 1-4: Representative yeast display plots for Rbg32.4 (left), Rbg32.5 (middle), and Rbg32.7 (right) at 1 μM h-IL-2Rβγ heterodimer.

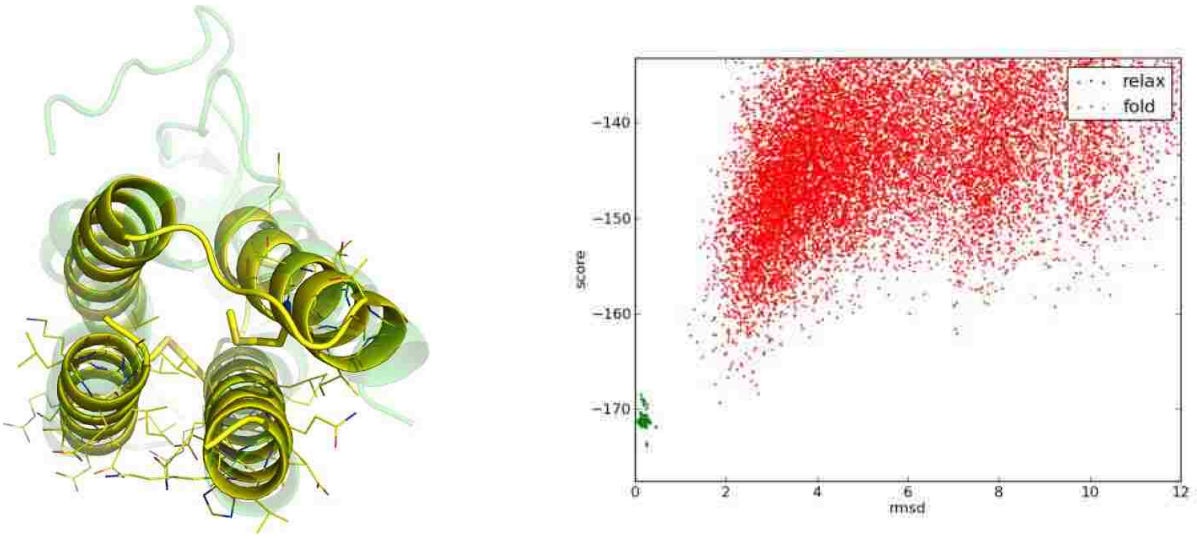


Figure 1-5: *Left*—Rbg40 (yellow) consists of 87 residues and four helices and closely recapitulates the topology of wild-type human IL-2 (green), especially at helices A, C, and D. Helix B on IL-2 (upper left helix) deviates slightly in Rbg40, as the adjacent loops connecting A to B and C to D (top) are excised. Other four-helix *de novo* designs closely resembled Rbg40 in structure. Computational model visualized using PyMOL.[60] *Right*—Forward folding experiments suggest a reasonable energetic minimum close to that of the intended structure.

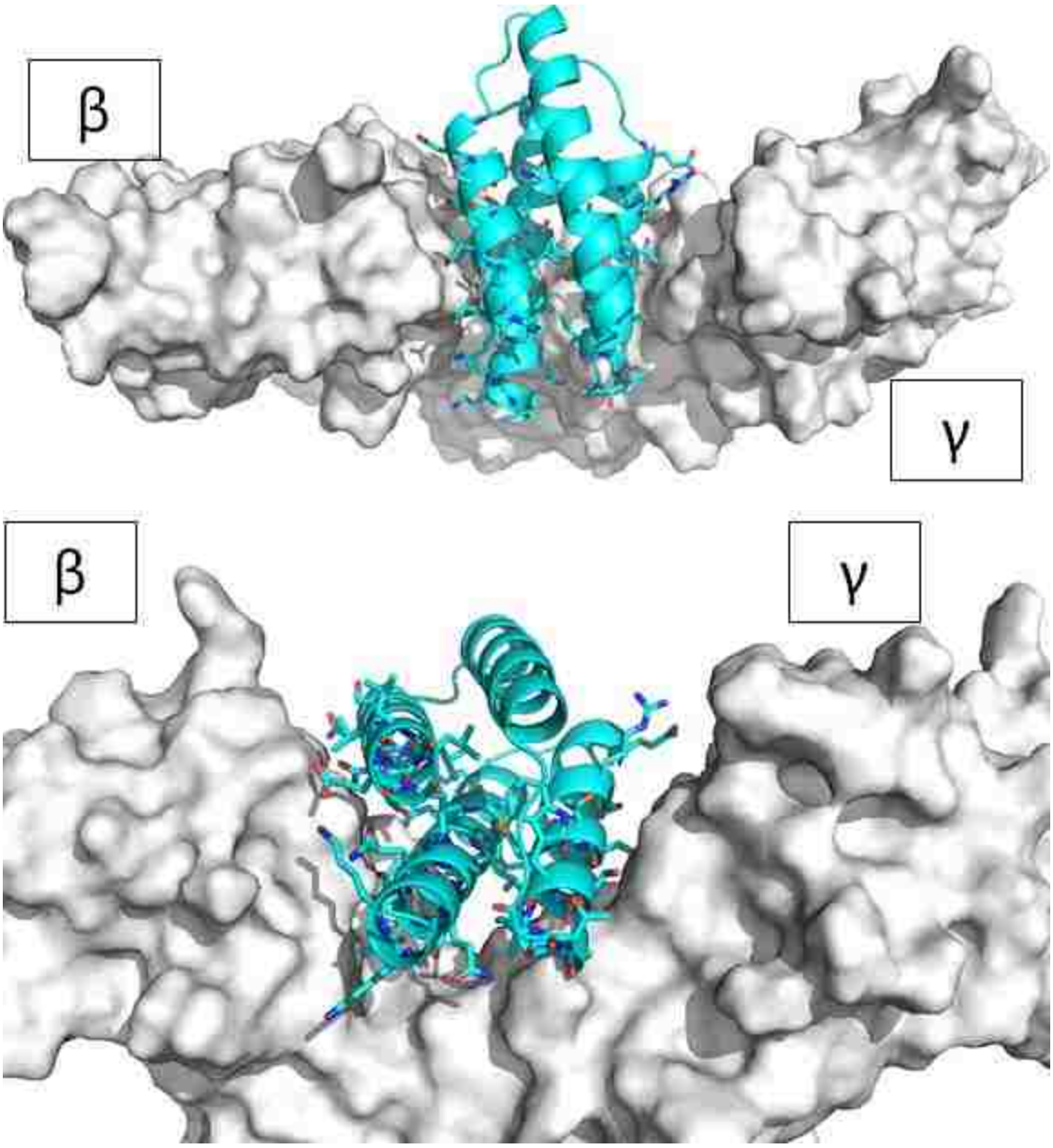


Figure 1-6: Computational model of Rbg40, a representative idealized design. *Top*—Top view of Rbg40 (cyan), a four-helix design which lies at the interface of the IL-2Rβ/γ_c receptor. *Bottom*—Side view of the same. Computational models visualized using PyMOL.[60]

Rbg37 Rbg38 Rbg39

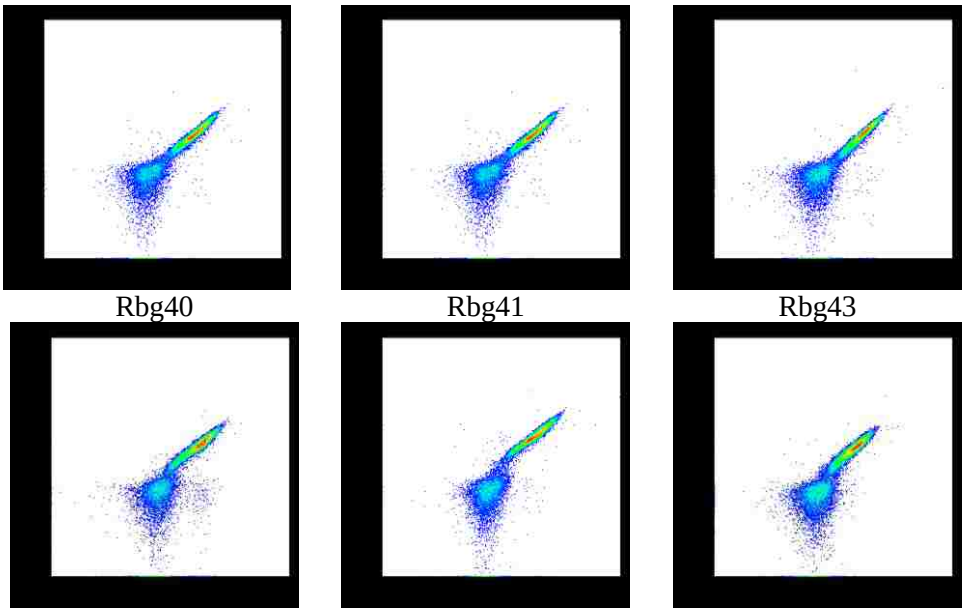


Figure 1-7: Yeast display screens for four-helix *de novo* designs Rbg37 (top left), Rbg38 (top middle), Rbg39 (top right), Rbg40 (bottom left), Rbg41 (bottom middle), and Rbg43 (bottom right). Designs were expressed on yeast cell surface, incubated with 1 nM b-hIL-2Rβγ, and labeled with SAPE. Rbg35, Rbg36, Rbg42, and Rbg44 also exhibited binding signal at higher concentrations of b-hIL-2Rβγ.

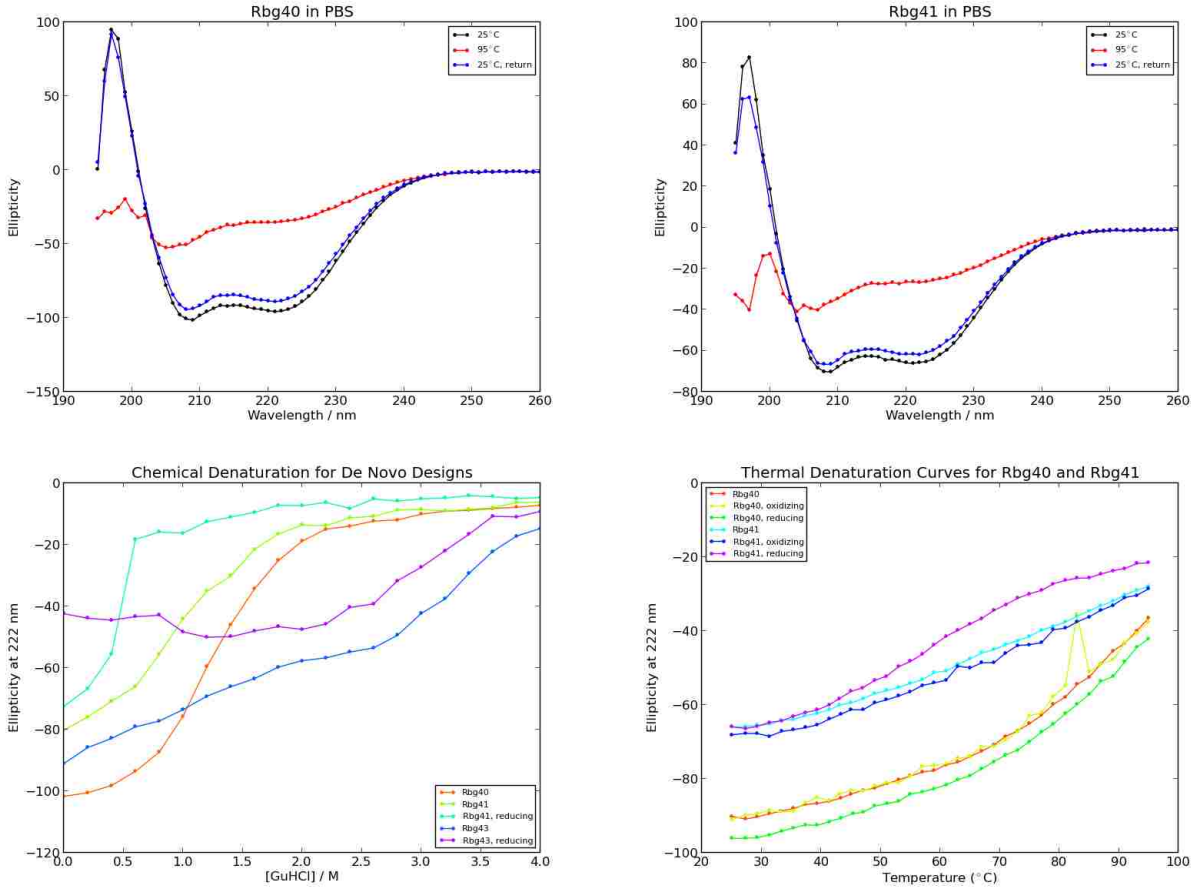
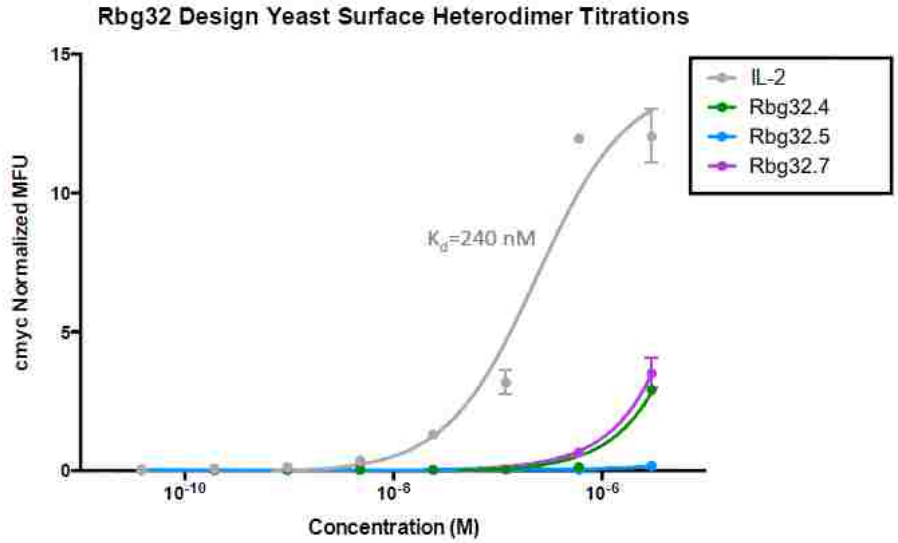
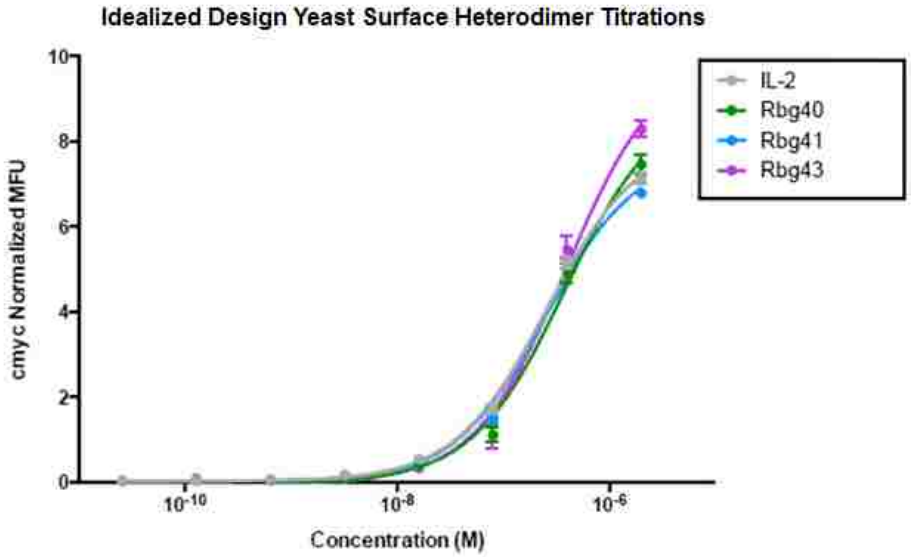


Figure 1-8: Circular dichroism experiments for Rbg40, Rbg41, and Rbg43. *Top left*—far-ultraviolet circular dichroism spectra for Rbg40 at 25°C (black), heated to 95°C (red), and recooled to 25°C (blue). *Top right*— far-ultraviolet circular dichroism spectra for Rbg41 at 25°C (black), heated to 95°C (red), and recooled to 25°C (blue). Both Rbg40 and Rbg41 retain strong alpha helical character after heating and recooling; Rbg43 (data not shown) exhibited greater loss of conformation. *Bottom left*—Chemical denaturation curves for Rbg40, Rbg41, and Rbg43. Rbg40 (red) retains structure at high concentrations of denaturant and exhibits cooperative folding. Denaturation experiments were also conducted for Rbg41 and Rbg43 under reducing conditions to disrupt their disulfide bonds. *Bottom right*—Thermal denaturation curves for Rbg40 and Rbg41, under normal, oxidizing, and reducing conditions. Neither melted completely under these conditions.

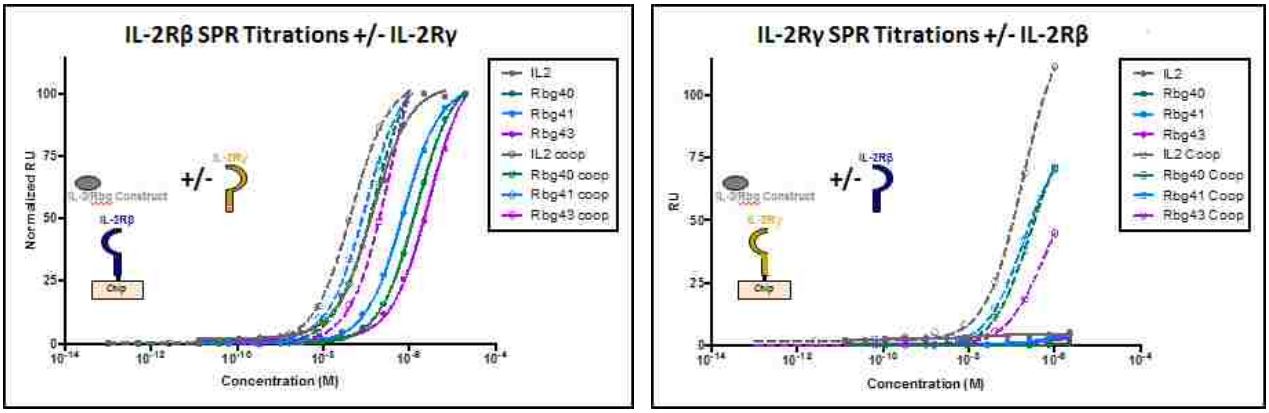


Binder	Kd / nM
IL-2	240
Rbg32.4	>10000
Rbg32.5	No Binding
Rbg32.7	>10000



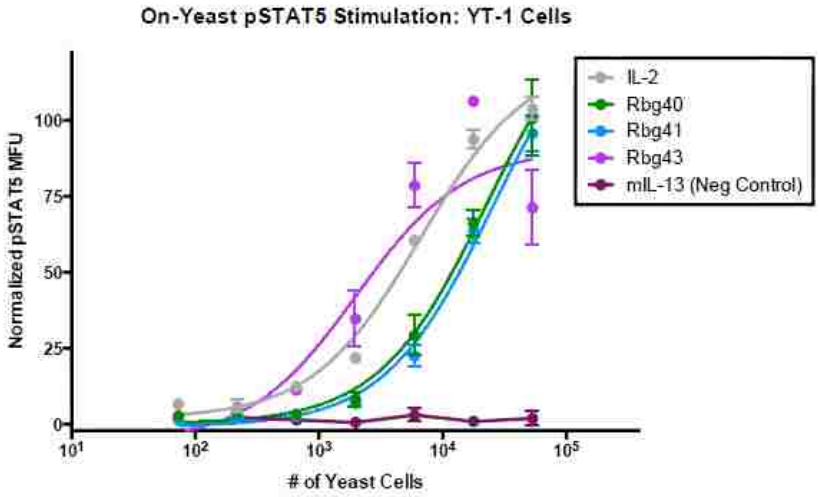
Binder	Kd / nM
IL-2	260
Rbg40	370
Rbg41	250
Rbg43	390

Figure 1-9: Yeast surface heterodimer titrations for three-helix MotifGraft designs (left) and four-helix idealized designs (right). First-generation three-helix MotifGraft binders exhibited binding affinities significantly weaker than that of wild-type IL-2 for b-hIL-2Rβγ, while first-generation four-helix idealized binders exhibited comparable affinities.[51]

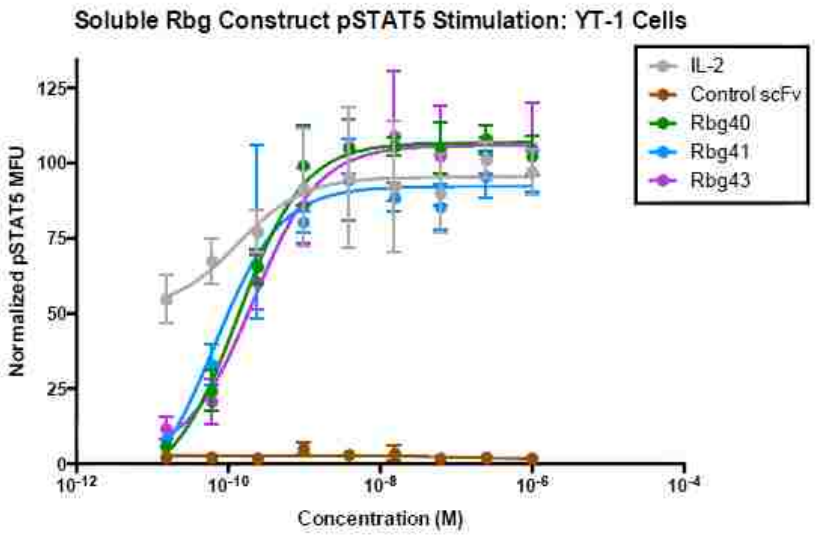


Analyte	IL-2Rβ	IL-2Rβ (cooperative)	γ_c	γ_c (cooperative)
IL-2	140	45	No Binding	180
Rbg40	1500	158	No Binding	260
Rbg41	720	94	No Binding	190
Rbg43	2900	34	No Binding	530

Figure 1-10: Surface plasmon resonance cooperativity studies. *Top left*—IL-2, Rbg40, Rbg41, and Rbg43 were flowed over immobilized IL-2Rβ in the absence or presence (cooperative) of γ_c. *Top right*—IL-2, Rbg40, Rbg41, and Rbg43 were flowed over immobilized γ_c in the absence or presence (cooperative) of IL-2Rβ. *Bottom*—measured binding affinities for each analyte in nanomolars. First-generation four-helix mimetics have reduced affinity for IL-2Rβ compared to IL-2 but comparable affinity for the intermediate receptor complex due to increased cooperativity.[51]



Binder	EC_{50} (# cells)
IL-2	6600
Rbg40	23000
Rbg41	25000
Rbg43	2000



Binder	EC_{50} / pM
IL-2	130
Rbg40	130
Rbg41	66
Rbg43	210

Figure 1-11: Stimulation assays for four-helix binders. In IL-2Rβγ-expressing YT-1 natural killer cell lines, Rbg40, Rbg41, and Rbg43 were able to stimulate STAT5 phosphorylation at comparable EC_{50} levels to that of IL-2 when presented on yeast surface (top) or presented in soluble form (bottom).[51]

Optimization of Designed Binders Against Human Interleukin-2 Receptor

Background

First-generation interleukin-2 (IL-2) mimetics Rbg32.7 and Rbg40 are two computationally designed proteins that bind biotinylated human IL-2Rβγ heterodimer, with Rbg40 achieving nearly comparable affinity to that of wild-type IL-2. Moreover, Rbg40 has also been demonstrated in STAT5 phosphorylation assays to elicit stimulation of natural killer cells expressing IL-2Rβγ with potency comparable to that of wild-type IL-2. However, in applications where immunostimulation is desirable, higher binding affinities may be necessary. IL-2 superkines have been reported to bind the IL-2Rβ/γ_c receptor form with significantly greater affinities than that of wild-type IL-2.[45]

Previously, site saturation mutagenesis had been used with deep sequencing to obtain comprehensive sequence-function maps on computationally designed inhibitors of influenza hemagglutinin.[31] By combining single mutations identified as beneficial, Whitehead *et al.* were able to obtain variants with subnanomolar binding affinity. These methods are applied to the work described in this section to obtain complete sequence-function maps for every residue on Rbg32.7 and Rbg40, and the resulting data are used to create combinatorial libraries of diversity exceeding 10^7, subsequently subjected to conventional directed evolution methods to isolate optimized binders against IL-2Rβ/γ_c.

Results

Sequence-Function Maps by Site-Saturation Mutagenesis

In order to optimize the "first-generation" three-helix MotifGraft and four-helix idealized binders for improved binding, site saturation mutagenesis (SSM) experiments were conducted to

provide sequence-function maps ("heatmaps") to better inform subsequent directed evolution. Rbg32.7 and Rbg40 were used as representative templates for MotifGraft and idealized designs, respectively.

All 74 residue positions were presented in the SSM library constructed for Rbg32.7, with all 1480 single mutants (including mutations to stop codon) represented between 18 to 1845 times in DNA reads. The heatmap for Rbg32.7 on the final round of selection was characterized by depletion at most residue positions (Figure 2-1), with many residues lying on the N-terminal helix from Val-4 to Glu-23 corresponding to those of low conservation (Figure 2-2). Because the N-terminal helix was not represented in the input motif for the parent Rbg21 nor subjected to targeted mutations in the focused library lib_Rbg21, it is less optimized on Rbg32.7 than other regions of the protein.

Specific strongly enriched mutations in SSM data include L9N and M5E, which suggests that the placement of the N-terminal helix in Rbg32.7 is closer to that of helix C of IL-2, as both mutations suggest the formation of hydrogen bonds that are present between the interaction between wild-type IL-2/IL-2Rβ (Figure 2-3). In addition, alanine residues Ala-14, Ala-36, and Ala-63 on the original Rbg32.7 sequence—lying on the three different helices, each pointing into the core—were replaced by larger hydrophobic residues in subsequent rounds of selection.

All 87 residue positions were represented in the SSM library constructed for Rbg40, with all 1740 single mutants represented between 1 to 468 times in DNA reads. In contrast to the observations for Rbg32.7, the SSM heatmap for Rbg40 (Figure 2-4) contains more positions of enrichment over wild-type but fewer mutations with extremely high enrichment ratios. Notably, mutations of Ser-42 to several hydrophobics (Phe, Leu, Val) presumably improve hydrophobic packing near the γ_c interface (Figure 2-5), to sit over the saddle formed by Tyr-103, Cys-160, and

Cys-209 of γ_c. In addition, mutations of Arg-73 and Leu-75 to Asp or Glu seem to confer better electrostatic complementary near the IL-2Rβ interface.

Affinity Maturation of Combinatorial Libraries

Combinatorial libraries for Rbg32.7 (Figure 2-6) and Rbg40 (Figure 2-8) were created based on SSM enrichment data, with amino acid diversities of 3.15×10^6 and 1.47×10^6, respectively. Rbg32.7 and Rbg40 combinatorial libraries were each subjected to four rounds of selection under increasingly stringent conditions according to Figure S2-1. Deep sequencing results on the final round of selection on Rbg32.7 combinatorial library included 594834 DNA reads in which all residues were consistent with the designed Rbg32.7 combinatorial library, with the eight most prevalent sequences by frequency accounting for 10.72% of these DNA reads. In addition, another 53741 DNA reads were identified as differing from the designed Rbg32.7 combinatorial library by exactly one amino acid substitution, with the four most prevalent such sequences by frequency accounting for 3.15% of these DNA reads. These twelve sequences were collectively assigned as second-generation three-helix MotifGraft binders with labels Rbg32.8A to Rbg32.8L (Figure 2-6).

The Rbg32.7 mutations identified as most enriched on SSM were also strongly enriched in combinatorial library selection. M5E and A68F were two of the most enriched mutations in SSM and were present in nearly all selected sequences of the combinatorial library, including all 12 second-generation binders. (L9N was the single most enriched mutation in SSM and was present in all sequences of the combinatorial library by design.) However, the most prevalent sequences that emerged were not simply a collection of the most enriched single point mutations, as covariation was evident in several trends that emerged, generally in the form of two or more

sterically incompatible mutations which would each individually improve metrics such as hydrophobic packing in the absence of others.

Likewise, deep sequencing results on the final round of selection on Rbg40 combinatorial library included 63856 DNA reads in which all residues were consistent with the designed Rbg40 combinatorial library, with the six most prevalent sequences by frequency accounting for 2.84% of these DNA reads. In addition, another 8049 DNA reads were identified as differing from the designed Rbg40 combinatorial library by exactly one amino acid substitution, with the four most prevalent such sequences by frequency accounting for 2.42% of these DNA reads. These twelve sequences were collectively identified as second-generation three-helix idealized binders, Rbg40.1A to Rbg40.1H (Figure 2-8). Interestingly, one of the strongest selective pressures manifested as a preference for Glu on position 62, a mutation for which enrichment was not evident on all SSM heatmaps. In retrospect, such a mutation results in better electrostatic complementarity and is consistent with the analogous Glu-67 in IL-2. The presence of R62E rendered mutations such as L75D or L75E less necessary for electrostatic complementarity, and in many sequences, L75Y (an enriched mutation with less significant electrostatic contributions) emerged instead.

The A68F mutation that was present on nearly all (99.70%) sequences of the final Rbg32.7 combinatorial library is particularly interesting, as it is analogous to the S42F (50.08%) or the similar S42Y (35.15%) mutations that were highly enriched in the Rbg40 library. In both cases, these substitutions represent mutations that were identified using directed evolution, rather than reversions to wild-type IL-2 where Ser-130 was present instead.

Yeast Surface Heterodimer Titrations

Optimized, second-generation Rbg32.7 (Figure 2-10) and Rbg40 (Figure 2-11) were displayed on yeast cell surface for titration against b-hIL-2Rβγ and exhibited significantly improved binding affinities compared to IL-2 control or wild-type first-generation Rbg32.7 and Rbg40. Of the twelve Rbg32.7 second-generation clones, all eleven which expressed had higher binding affinities for b-hIL-2Rβγ (K_d = 7.2 nM for Rbg32.7H) than did IL-2, representing a roughly 1000-fold improvement in binding affinity compared to first generation Rbg32.7. All ten Rbg40 second-generation clones also had higher binding affinities for b-hIL-2Rβγ (K_d = 1.4 nM for Rbg40.1F). Interestingly, the strongest binders within these second-generation clones were not simply the most prevalent ones selected from combinatorial library sorting (Rbg32.7A and Rbg40.1A).

Methods

Creation of SSM libraries

Forward primers and reverse primers were designed for each amino acid residue on Rbg32.7 (Table S2-1) and Rbg40 (Table S2-2), resulting in a "left" PCR product with a degenerate NNK codon and a "right" PCR product when amplified with COR and COF primers, respectively. Amplification of "left" and "right" products by COF and COR primers yielded a series of template products each consisting of a degenerate NNK codon at a different residue position. These products were pooled to yield the SSM library. SSM libraries were transformed by electroporation into conditioned *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector, using the protocol previously described by Benatuil *et al*.[57]

Creation of combinatorial libraries

Combinatorial libraries for Rbg32.7 and Rbg40 against hIL-2Rβ/γ_c were created by assembly PCR using oligonucleotides shown in Table S2-3. Genes were transformed by electroporation into conditioned *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector, using the protocol previously described by Benatuil *et al*.[57]

Fluorescence-activated cell sorting

Fluorescence-activated cell sorting was conducted using human IL-2Rβ and γ_c receptors, expressed as acid/base zipper heterodimers with C-terminal BAP tag (Table S1-4) to facilitate biotinylation (b-hIL-2Rβγ), provided as a gift from the Garcia Lab. Yeast cells containing SSM or combinatorial libraries were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm and induced in 1 mL SGCAA for expression at 30°C for 16 h. Five million (5×10^6) cells were collected; incubated with b-hIL-2Rβγ for 2 h in phosphate buffered saline supplemented with 0.5% BSA and 2 mM EDTA (PBE); and labeled with 5 μg streptavidin, R-phycoerythrin conjugate (SAPE, Immunology Consultants Laboratory) and 5 μg FITC conjugated chicken anti-C-Myc (anti-C-Myc-FITC, Life Technologies) in 250 μL PBE for 15 min on ice; and collected in 1 mL c-Trp-Ura media based on a combination of gates based on FSC, SSC, PE, and FITC signals. Selection process was repeated for three or four rounds with concentration of b-hIL-2Rβγ and incubation temperature varying as indicated in Figure S2-1. For rounds of selections involving a dissociation step, cells were incubated with 1 mL PBE for 1 h at 37°C after incubation with b-hIL-2Rβγ but before labeling with SAPE or anti-C-Myc-FITC. FACS experiments were performed on the BD Influx cell sorter from BD Biosciences and on the SH800 Cell Sorter from Sony Biotechnology.

Yeast-display titrations

Yeast cells were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm, and induced in 1 mL SGCAA for expression at 30°C for 16 h. Fifty thousand (5×10^5) cells were collected; incubated with b-hIL-2Rβγ for 2 h at 37°C in PBE; and labeled with 1 μg streptavidin, Alexa Fluor 647 conjugate (SA-647) and 1 μg anti-C-Myc-FITC in 50 μL PBE for 15 min on ice. Human IL-2 displayed on EBY100 strain cells containing pCT302 vector[45] served as controls where applicable. Measurements were performed on Accuri C6 flow cytometer. Cells were gated using FlowJo v10 software and median FL-4 values were used for titration data. Data were fitted using Prism 6 software, using a non-linear one site total saturation binding model.

Library preparation and sequencing

DNA from SSM or combinatorial libraries were prepped for sequencing using a protocol adapted from Chevalier *et al*. Briefly, aliquots of culture from naïve and selected libraries were grown in 1 mL c-Trp-Ura media at 37°C overnight or to an optical density of 4.0 at 600 nm. Cells were treated with Zymolyase to lyse their cell walls (Zymo research), subjected to Miniprep to isolate plasmid DNA (Qiagen), and purified of residual genomic or ssDNA with Exonuclease I and Lambda Exonuclease (New England Biolabs). Two PCR steps were then included to add Illumina adapter regions and to introduce unique barcodes for each library. Sequencing was performed using the MiSeq desktop sequencer.

Sequence-Function Maps

For naïve and selected libraries, forward and reverse paired-end FASTQ files from MiSeq were merged with PEAR[61] using default p-value, minimum overlap, assembly length, and quality score threshold options. For each sequence in the merged FASTQ file, the sub-sequence

between restriction sites CATATG and CTCGAG was identified as the gene, checked for appropriate length compared to the SSM template, and translated into the corresponding amino acid sequence. The number of occurrences of each amino acid sequence was tabulated, with only wild-type and single-point mutations included in the final analysis. Enrichment values were determined as follows:

$$E_{s,N} = \log_2\left(\frac{p_{s,N}/p_{s,0}}{p_{wt,N}/p_{wt,0}}\right)$$

$$p_{i=s,j} = \frac{c_{s,j} + k}{\sum_i c_{i,j} + k}$$

where the enrichment value $E_{s,N}$ of a particular sequence in round N was given by the binary logarithm of the ratio of $p_{s,N}$, the proportion of sequence in round N, and $p_{s,0}$, the proportion of sequence in the naïve library, normalized by the corresponding ratio for the proportion of wild-type template sequence $p_{wt,N}$ in round N over its proportion in the naïve library. Proportion values $p_{s,j}$ were calculated by taking the ratio between the number of observations of sequence s in round j, plus a pseudocount factor $k = 0.1$, over the number of observations of all wild-type and single mutants, plus a pseudocount factor $k = 0.1$.

Recombinant expression

Plasmids containing genes in pETCON yeast expression vector were extracted from yeast EBY100 cells using Zymolyase from Zymo Research. Genes of interest were amplified from pETCON with primers shown in Table S2-4, inserted into pET29b *E. coli* expression vector using Gibson assembly[58], and transformed into *E. coli* XL10-gold strain cells by heat shock. Plasmids containing genes encoding designs of interest in pET29b vector were extracted from *E. coli* XL10-gold cells by Miniprep (Qiagen) and transformed into *E. coli* BL21(DE3) strain cells.

Cells were grown at 37°C in 0.5 L Terrific Broth to an optical density of 0.6 at 600 nm and induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for expression at 18°C for 6 h. Cell pellets were harvested by centrifugation and lysed by sonication. Protein was purified by Ni-NTA chromatography (eluted in PBS with 250 mM imidazole), verified by SDS-PAGE and mass spectrometry, and again purified for monomers via size exclusion chromatography (eluted in PBS).

Discussion

SSM has provided useful information on sequence-function relationships, serving as a way to not only identify potentially beneficial mutations, but also in the absence of high-resolution crystallographic data, as a means to partially validate the structure of the computational designs. Affinity maturation methods that combine SSM data with traditional directed evolution techniques on combinatorial libraries can yield significant improvement in binding affinity (roughly 1000-fold for Rbg32.7 and 100-fold for Rbg40).

While SSM provides valuable information on mutations that are beneficial for binding, it is worth noting that the best sequences that emerged were not simply collections of the single best point mutations. Epistasis was most evident in the case of two or more sterically incompatible mutations which would each individually improve metrics such as hydrophobic packing. For instance, SSM data suggest that Ala-14 and Val-18 of Rbg32.7 would benefit from mutations to larger, hydrophobic residues, as both are positioned on consecutive turns of a helix that points into a relatively poorly packed core. However, two such mutations are less compatible simultaneously because of the possibility of steric clashes or the presence of too many hydrophobic residues compromising the overall stability of the desired conformation. As a result, whereas position 14 contains 29.72% Phe and 47.55% Tyr in the final Rbg32.7 combinatorial

library, those frequencies drop to 26.37% and 25.06%, respectively, when position 18 is occupied by Trp. Likewise, Ala-14, Ala-36, and Ala-63 lie on the three different helices of Rbg32.7 and are oriented towards the interior. While mutations of these residues to larger hydrophobic ones appear enriched on SSM, they are not simultaneously compatible. When position 36 is occupied by Ala, position 36 contains only 11.86% Ala; when Trp-36 is present instead, position 63 contains 53.15% Ala. (Note that Ala-14 is excluded from the designed library, so potential steric issues selective pressures largely exert at positions 36 and 63 instead.)

It also appears that proteins which are less optimized for their intended target yield more conclusive SSM data. In the case of the relatively unoptimized Rbg32.7, the heatmap was dominated by several mutations with extremely high enrichment ratios (M5E, L9N, A68F) which persisted through multiple rounds of sorting on the combinatorial library and appeared to contribute to the drastically improved binding affinity of the second-generation three-helix variants. In contrast, the heatmap for Rbg40 featured more mutations that were enriched over wild-type but very few specific mutations that were extremely highly enriched; rather, positions such as Gln-8, Arg-22, Ser-23, Ser-42, and Ala-65 were permissive towards mutations to a wide range of hydrophobic residues.

Tables and Figures

Design	**Sequence**
Rbg32.8A	SAQVELEDNARELYIACGKALKEGKVEDACTYCEHALLDLQQLQVLYPESVARTDYEQMITEVQRRI FNCEKVL
Rbg32.8B	SAQVELEDNARELLIACLKADKEGKVEDACTYCEHALLDLQQLIVLYPESVARTDYEQMITETQRRI FNCEKVL
Rbg32.8C	SAQVELEDNARELYIACVKAVKVGKVEDACTFCEHALLDLQQLKVLYPESVARTDYEQMITELQRRI FNCEKVL
Rbg32.8D	SAQVELEDNARELFIACGKAVKVGKVEDACTFCEHALLDLQQLIVLYPESVARTAYEQMITEAQRRI FNCEKVL
Rbg32.8E	SAQVELEDNARELYIACLKAVKVGKVEDACTYCEHGLLDLQQLIVLYPESVARTDYEQMITEMQRRI FNCEKVL
Rbg32.8F	SAQVELEDNARELYIACLKADKEGKVEDACTICEHALLDLQQILVLYPESVARTAYEQMITEMQRRI FNCEKVL

Rbg32.8G	SAQVELEDNARELHIACWKADKEGKVEDACTFCEHALLDLQQLIVLYPESVARTDYEQMITETQRRI FNCEKVL
Rbg32.8H	SAQVELEDNARELYIACVKADKVGKVEDACTFCEHSLLDLQQLLVLYPESVARTAYEQMITELQRRI FNCEKVL
Rbg32.8I	SAQVELEDNARELLIACGKAVKEGKVEDACTYCEHALLDLQQLKVLYPESVARTDYEQMITEMQRRI FNCEKFL
Rbg32.8J	SAQVELEDNARELFIACVKADKVGKVEDACTICEHALLDLQQLLVLYPESVARTDYEQMITELQRRI FKCEKVL
Rbg32.8K	SAQVELEDNARELYIACVKAHKVGKVENACTYCEHALLDLQQLLVLYPESVARTDYEQMITEMQRRI FNCEKVL
Rbg32.8L	SAQVELEDNARELYIACLEAVKEGKVEDACTYCEHALLDLQQLQVLYPESVARTDYEQMITEVQRRI FNCEKVL

Table 2-1: Second-generation three-helix MotifGraft binders against human IL-2Rβ/γ_c.

Design	**Sequence**
Rbg40.1A	STKKTQLLAEHALLDAFMMLNVVPEPNEKLNRIITTMQSWIYTGKIDADGAKELAKEVEELEQEYE KRGIDVEDDASNLKVILLELA
Rbg40.1B	STKKTQLLAEHALLDAHMMLNMLPEPNEKLNRIITTMQSWIHTGKIDGDGAQELAKEVEELEQEYE KRGIDVEDEASNLKVILLELA
Rbg40.1C	STKKTQLLAEHALLDAFMMLNMVPEPNEKLNRIITTMQSWIFTGKIDGDGAKELAKEVEELEQEFE KRGIDVEDEASNLKVILLELA
Rbg40.1D	STKKTQLLAEHALLDALMMLNMVPEPNEKLNRIITTMQSWIFTGKIDGDGAQELAKEVEELEQELE KRGIDVEDYASNLKVILLELA
Rbg40.1E	STKKTQLLAEHALLDAHMMLNVVPEPNEKLNRIITTMQSWIYTGKIDRDGAQELAKEVEELEQELE KRGIDVDDDASNLKVILLELA
Rbg40.1F	STKKTQLLAEHALLDALMMLNLLPEPNEKLNRIITTMQSWIFTGKIDGDGAQELAKEVEELEQEHE KRGIDVEDYASNLKVILLELA
Rbg40.1G	STKKTQLLAEHALLDAYMMLNMVPEPNEKLNRIITTMQSWILTGKIDSDGAQELAKEVEELEQELE KRGIDVDDDASNLKVILLELA
Rbg40.1H	STKKTHLLAEHALLDAYMMLNVMPEPNEKLNRIITTMQSWIFTGKIDGDGAKELAKEVEELEQEFE KRGIDVDDDASNLKVILLELA
Rbg40.1I	STKKTQLLAEHALLDAYMMLNLVPEPNEKLNRIITTMQSWIFTGKIDADGAQELAIEVEELEQEYE KRGIDVDDYASNLKVILLELA
Rbg40.1J	STKKTQLMAEHALLDAFMMLNVLPEPNEKLNRIITTMQSWIFTGKIDGDDAQELAKEVEELEQELE KRGIDVDDDASNLKVILLELA

Table 2-2: Second-generation four-helix idealized binders against human IL-2Rβ/γ_c.

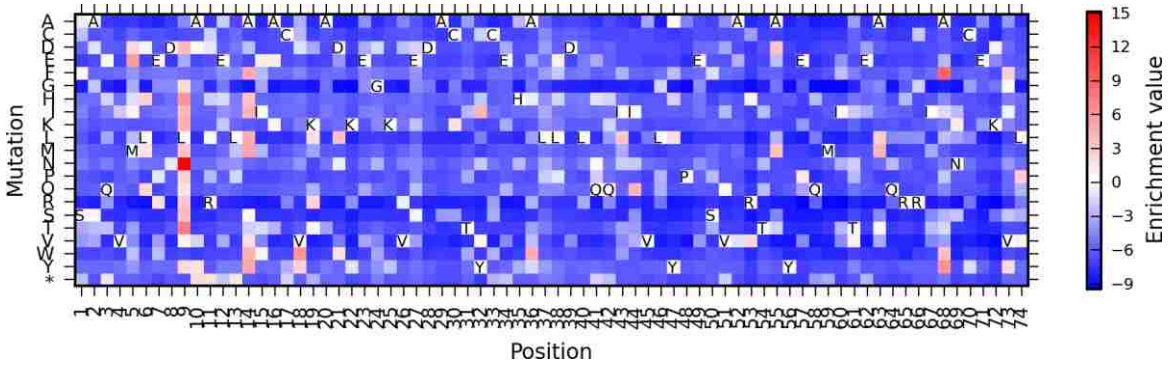


Figure 2-1: Enrichment data for Rbg32.7 SSM round 2 library. The second round of selection was performed binding against 50 nM b-hIL-2Rβγ at 4°C. Enrichment data for round 1 (4°C) not shown.

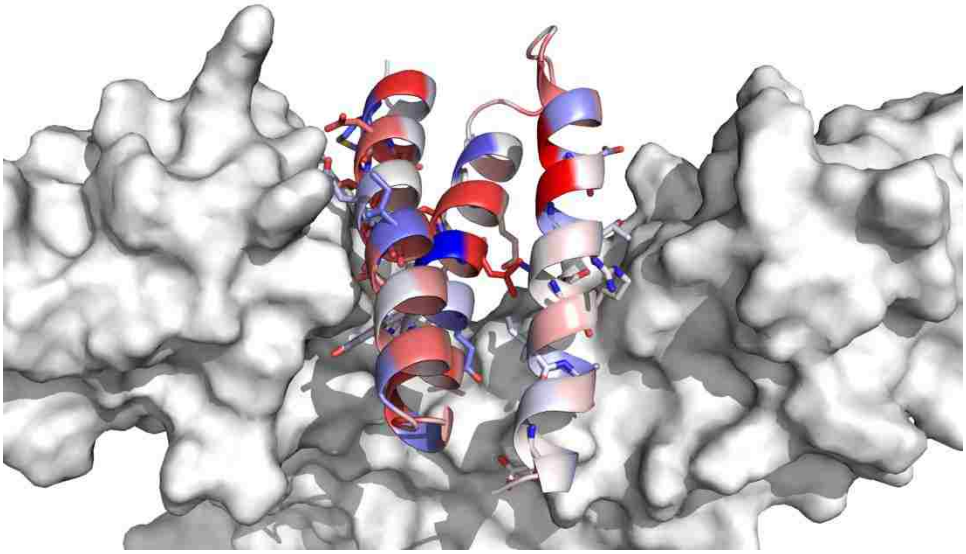


Figure 2-2: Structure of Rbg32.7 colored by conservation of original amino acid in SSM selection, with red indicating poor conservation and blue indicating high conservation. The left-most helix is the N-terminal helix which contains contact residues exclusively against IL-2Rβ. Because this helix was not included in the input motif during the design process and did not contain any mutations in the focused library lib_Rbg21, it is less optimized against b-hIL2Rβγ than the other two helices.

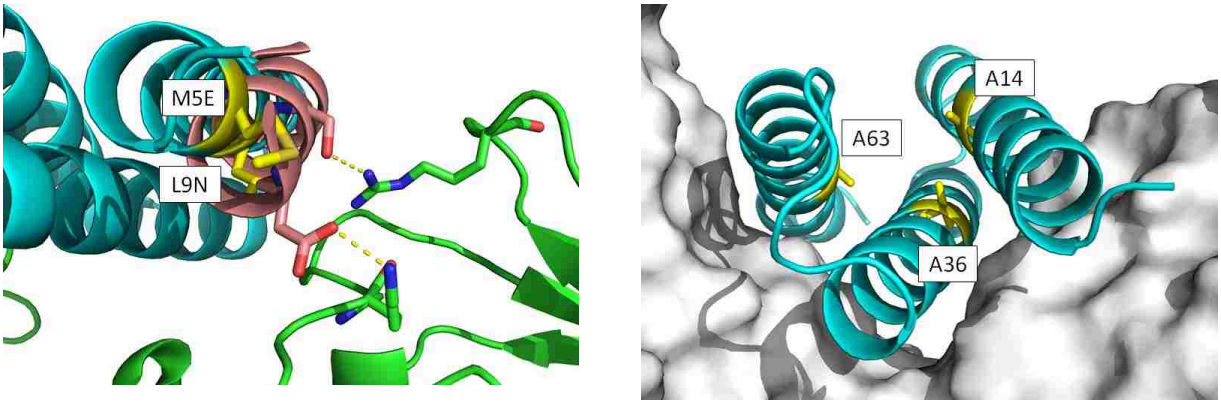


Figure 2-3: Selected mutations from Rbg32.7 SSM enrichment data. *Left*—enriched mutations on the N-terminal helix of Rbg32.7 (cyan), contacting IL-2Rβ (green), such as M5E and L9N (yellow), suggest that this helix probably resembles helix C of IL-2 (pink) in which Asp-84 and Ser-87 of IL-2 form hydrogen bonds with Lys-71 and Arg-42 of IL-2Rβ, respectively, more closely than suggested by the computational model. *Right*—enriched mutations A14, A36, and A63 (yellow)on different helices of Rbg32.7 (cyan) each point towards the hydrophobic core, suggesting that improved hydrophobic packing of the interior contributes to the enrichment of these mutations in the selected clones. Computational models visualized using PyMOL.[60]

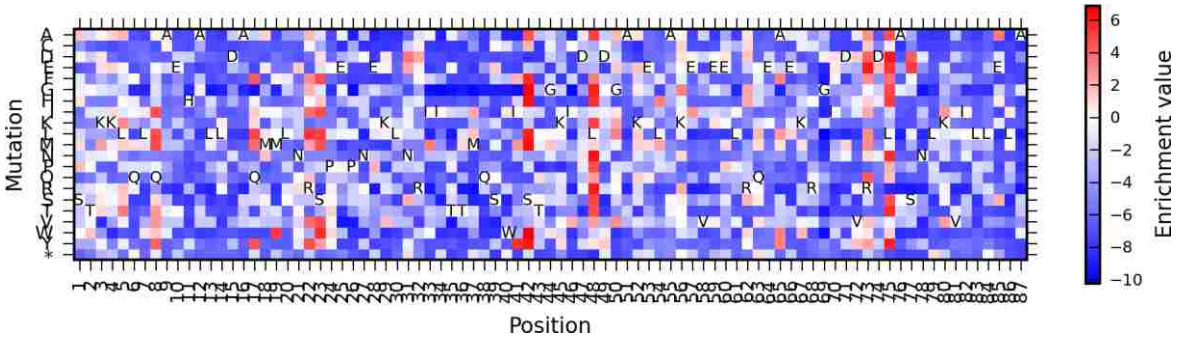


Figure 2-4: Enrichment data for Rbg40 SSM round 3 library. The selection for the heatmap above was performed binding against 200 pM b-hIL-2Rβγ at 37°C. Enrichment data for rounds 1 and round 2 and for other temperatures not presented.

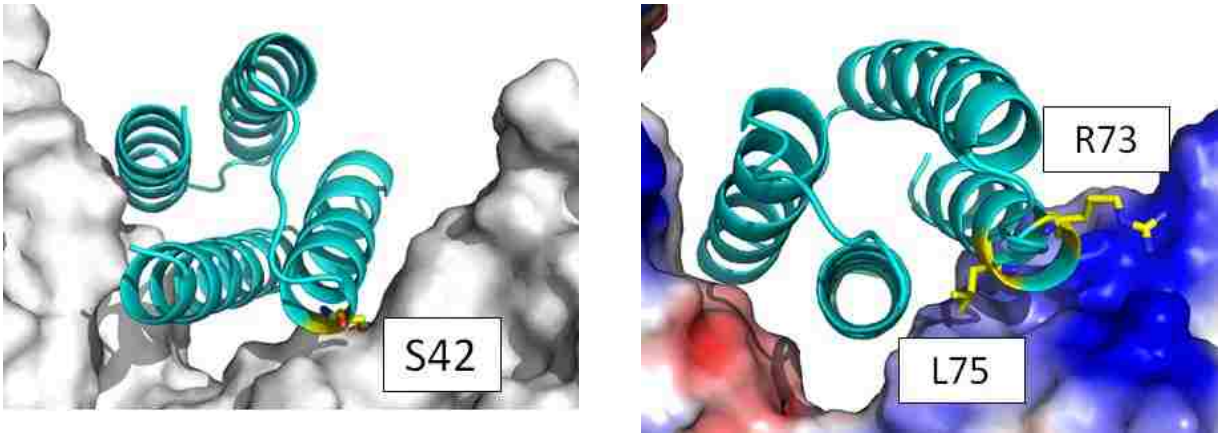
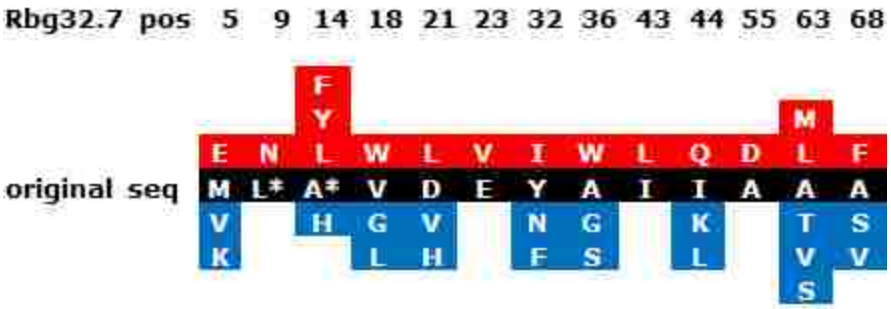


Figure 2-5: Selected mutations from Rbg40 SSM enrichment data. *Left*—mutation of Ser-42 (yellow) to Phe, Tyr, Leu allows the new residue to pack over the saddle formed by Tyr-103, Cys-160, and Cys-209 (magenta) on γ_c This highly enriched mutation is also observed on Rbg32.7 at Ala-68, where all selected sequences in the combinatorial library converge to Phe. *Right*—electrostatic surface representation of IL-2Rβ/γ_c reveals large patches of positive charge (blue) on IL-2Rβ near the interface with Rbg40 (cyan). Mutation of Leu-73 and Lys-75 (yellow) to Asp or Glu results in better electrostatic complementarity. Computational models visualized using PyMOL.[60]



Rbg32.8A	E	N	Y	G	L	E	Y	A	L	Q	D	V	F
Rbg32.8B	E	N	L	L	D	E	Y	A	L	I	D	T	F
Rbg32.8C	E	N	Y	V	V	V	F	A	L	K	D	L	F
Rbg32.8D	E	N	F	G	V	V	F	A	L	I	A	A	F
Rbg32.8E	E	N	Y	L	V	V	Y	G	L	I	D	M	F
Rbg32.8F	E	N	Y	L	D	E	I	A	I	L	A	M	F
Rbg32.8G	E	N	H	W	D	E	F	A	L	I	D	T	F
Rbg32.8H	E	N	Y	V	D	V	F	S	L	L	A	L	F
Rbg32.8I	E	N	L	G	V	E	Y	A	L	K	D	M	F
Rbg32.8J	E	N	F	V	D	V	I	A	L	L	D	L	F
Rbg32.8K	E	N	Y	V	H	V	Y	A	L	L	D	M	F
Rbg32.8L	E	N	Y	L	V	E	Y	A	L	Q	D	V	F

Figure 2-6: Second-generation three-helix MotifGraft binders against human IL-2Rβ/γ_c. Combinatorial library for Rbg32.7 contained 13 mutations from original Rbg32.7 with a combined diversity of 3.15×10^6. Shown in black background are the amino acid identities of the original sequence. Red background indicates desirable amino acids allowed by the designed codon; blue background indicates unintended amino acids allowed by the designed codon. Amino acids from the original sequence are allowed by the designed codon unless indicated by an asterisk. Identities of the twelve clones (Rbg32.8A to Rbg32.8L) at indicated positions are listed at the bottom. Sequences with exactly one amino acid substitution not present in the designed combinatorial library (Rbg32.8I—V73F, Rbg32.8J—N69K, Rbg32.8K—D28N, Rbg32.8L—K19E) are indicated in italics.

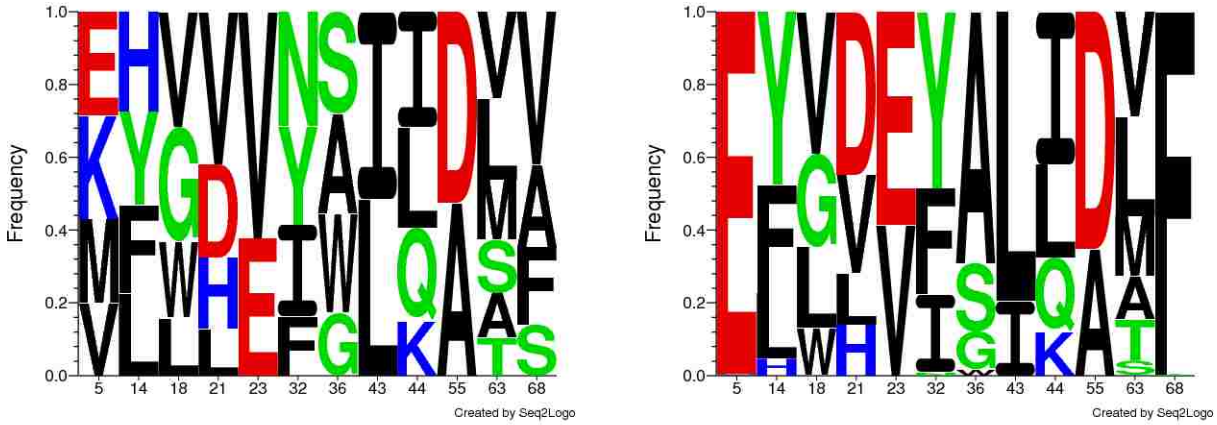


Figure 2-7: Sequence logos for Rbg32.7 combinatorial library, with the height of each letter corresponding to the frequency of the corresponding amino acid at that position. *Left*—sequence logo for naïve library depicting all allowable mutations present in roughly equal frequencies at each codon. *Right*—after four rounds of sorting, there is strong convergence towards Glu-5 and Phe-68, in addition to weaker selective pressures elsewhere. Sequence logos generated using Seq2Logo.[62]

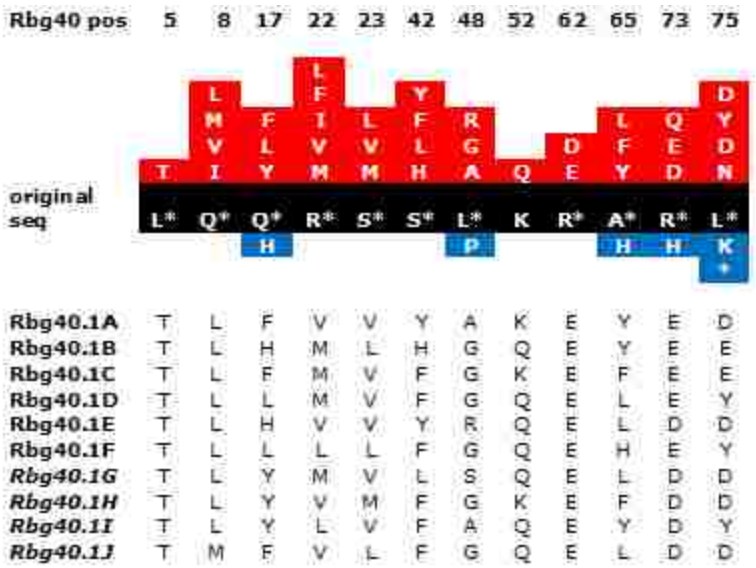


Figure 2-8: Second-generation four-helix idealized binders against human IL-2Rβ/γ_c. Combinatorial library for Rbg40 contained 12 mutations from original Rbg40 with a combined amino acid diversity of 1.47×10^6. Shown in black background are the amino acid identities of the original sequence. Red background indicates desirable amino acids allowed by the designed codon; blue background indicates unintended amino acids allowed by the designed codon. Amino acids from the original sequence are allowed by the designed codon unless indicated by an asterisk. Identities of the ten clones (Rbg40.1A to Rbg40.8J) at indicated positions are listed at the bottom. Sequences with exactly one amino acid substitution not present in the designed combinatorial library (Rbg40.1G—L48S, Rbg40.1H—Q6H, Rbg40.1I—K56I, Rbg40.1J—G50D) are indicated in italics.

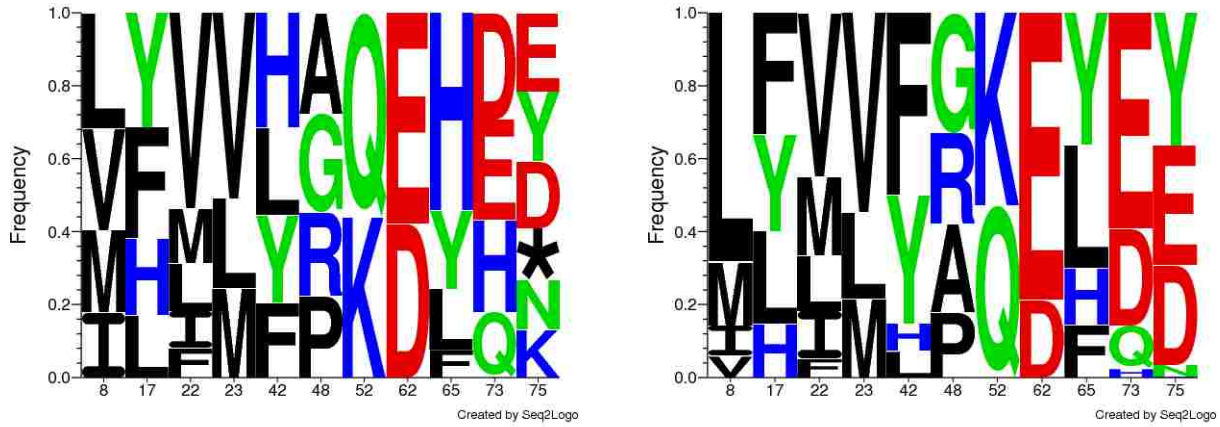
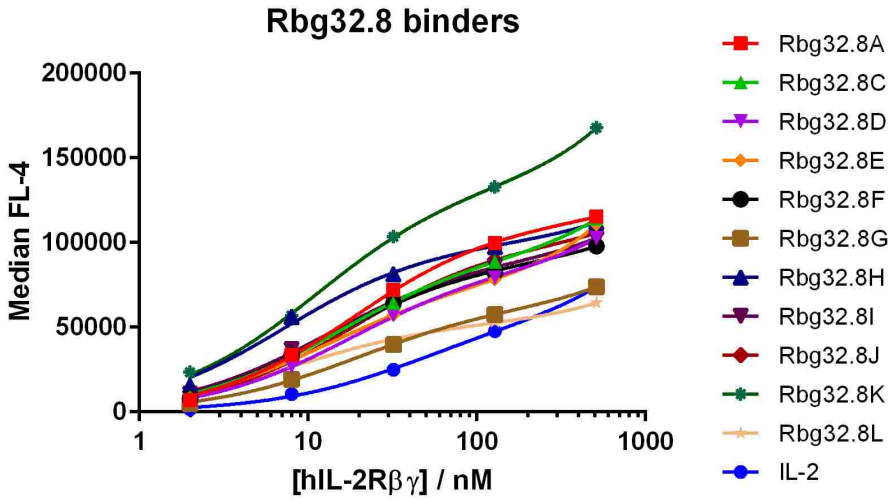
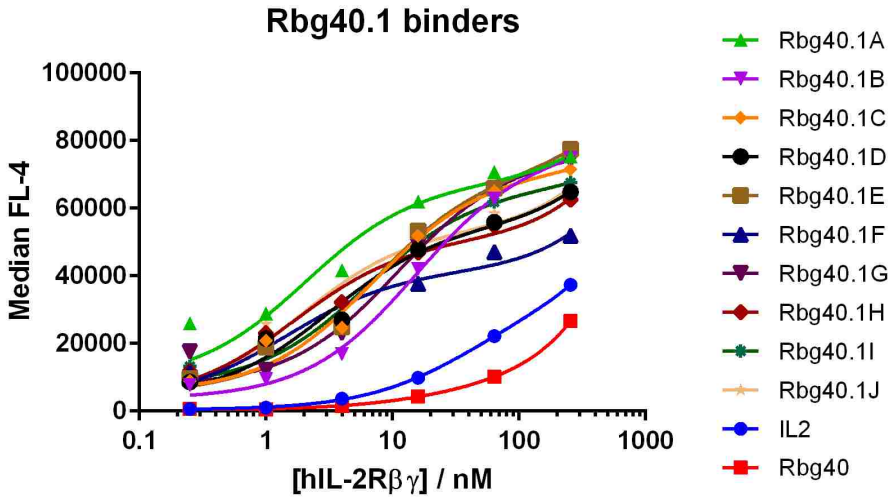


Figure 2-9: Sequence logos for Rbg40 combinatorial library, with the height of each letter corresponding to the frequency of the corresponding amino acid at that position. *Left*—sequence logo for naïve library depicting all allowable mutations present in roughly equal frequencies at each codon. *Right*—after four rounds of sorting, there is moderate selection for Leu-8 and Glu-62 and for Phe and Tyr at position 42. Sequence logos generated using Seq2Logo.[62]



Design	K_d / nM
Rbg32.8A	18.56
Rbg32.8C	14.4
Rbg32.8D	17.03
Rbg32.8E	12.01
Rbg32.8F	13.13
Rbg32.8G	18.6
Rbg32.8H	7.198
Rbg32.8I	12.5
Rbg32.8J	18.36
Rbg32.8K	10.88
Rbg32.8L	7.696
IL-2	41.41

Figure 2-10: Yeast surface titration curves for second-generation Rbg32.8 clones against hIL-2Rβ/γ_c with IL-2 control. Rbg32.8B did not express.



Design	K_d / nM
Rbg40.1A	2.055
Rbg40.1B	15.13
Rbg40.1C	7.013
Rbg40.1D	3.177
Rbg40.1E	6.876
Rbg40.1F	1.418
Rbg40.1G	11.15
Rbg40.1H	1.679
Rbg40.1I	6.219
Rbg40.1J	1.797
IL-2	38.28
Rbg40	22.87

Figure 2-11: Yeast surface titration curves for second-generation Rbg40.1 clones against hIL-2Rβ/γ_c with IL-2 and Rbg40 control.

Optimization of Designed Binders Against Mouse Interleukin-2 Receptor

Background

Sequence-function maps from site saturation mutagenesis have provided enrichment data to guide the optimization of computationally designed proteins, as illustrated in the maturation of influenza virus hemagglutinin.[31] Moreover, such data has been used to not only improve binding affinity against a target of interest, but to improve affinity and even modulate specificity for closely related targets such as H1 and H5 subtype hemagglutinin. When combined with structural information, the empirical data yielded from such experiments can help improve the force fields and design methods used in computational protein design.

In this section, empirical data from SSM experiments are used to engineer interleukin-2 mimetics with improved binding affinity to the mouse IL-2Rβ/γ_c receptor complex in the absence of high-resolution structural information on the IL-2Rβ/γ_c or the IL-2/IL-2Rβ/γ_c complexes. Furthermore, the development of such binders with improved affinity for mouse receptor would better enable *in vivo* studies on the biological effects of these mimetics in mouse animal models.

Results

Cross-Reactivity with Mouse Receptors

Human and mouse IL-2 cytokine and receptor subunits share modest sequence similarity (Table 3-1), allowing some attenuated cross-reactivity when cytokine of one species engages the receptor of another (Figure 3-1). However, first-generation four-helix designs Rbg40, Rbg41, and Rbg43 have reduced binding affinities for the individual mIL-2Rβ subunit and for the mIL-2Rβ/γ_c complex, despite exhibiting comparable binding affinity for the hIL-2Rβ/γ_c complex (Figure 1-10). Two four-helix variants (Table 3-2) were designed by comparing residues on

Rbg40.1F to analogous ones on hIL-2, then substituting based on a sequence comparison between mIL-2 and h-IL2, but the amino acid substitutions introduced did not improve the binding affinities of these variants for mIL-2Rβγ above that of Rbg40.1F (Figure 3-2).

Sequence-Function Maps

In order to optimize Rbg40.1F for improved binding against mIL-2Rβ/γ_c, site saturation mutagenesis (SSM) experiments were conducted to provide sequence-function maps to better inform subsequent rounds of directed evolution. All 87 residue positions were presented in the SSM library constructed for Rbg40.1F, with all 1740 single mutants (including mutations to stop codon) represented between 9 to 9809 times in DNA reads.

In the absence of high-resolution structure for mouse IL-2Rβ/γ_c or the IL-2/IL-2Rβ/γ_c complexes, the empirical enrichment data obtained from such experiments against biotinylated mouse IL-2Rβ/γ_c (b-mIL-2Rβγ) alone cannot be validated with computational predictions or human intuition that rely on accurate structural information. Therefore, the Rbg40.1F SSM library was selected against b-mIL-2Rβγ and b-hIL-2Rβγ in parallel so that known structural information about the latter could allow for inferences on the nature of the enriched mutations in selections against the former.

Sequence-function maps for Rbg40.1F when selected against mIL-2Rβ/γ_c ("mouse heatmap") and hIL-2Rβ/γ_c ("human heatmap") are shown in Figure 3-3 and Figure 3-4, respectively. Most of the enriched residues on the mouse or human heatmaps were included in the corresponding combinatorial libraries as shown in Figure 3-5. Whereas many of the enriched mutations that appear on the mouse heatmap are located at the interface, the enriched mutations that emerged on the human heatmap are distributed more on surface residues and loops. Furthermore, the mouse heatmap featured several highly enriched mutations with enrichment

ratios (~2^9) significantly higher than those in the human heatmap (~2^6). Both observations are consistent with the expectation that the mouse heatmap contain a greater proportion of interface residues with high enrichment ratios, as the starting template Rbg40.1F had previously been already optimized against hIL-2Rβ/γ_c.

Interestingly, none of the substitutions introduced by sequence-based substitutions were particularly enriched in the mouse heatmap. One residue of note was Ala-9, as both A9F and A9M were highly enriched mutations on the human heatmap which made intuitive sense structurally given that the replacement of Ala-9 with a larger hydrophobic residue would better pack the protein core (Figure 3-6). Because these mutations affect the stability of the protein, rather than the binding interface, one might expect them to be also enriched on the mouse heatmap, but A9F and A9M are depleted in the mouse SSM libraries instead.

Affinity Maturation of Combinatorial Libraries

Combinatorial libraries for mouse (Figure 3-7) and for human (Figure 3-8) were created based on Rbg40.1F SSM enrichment data, with amino acid diversities of 1.05×10^6 and 3.15×10^6, respectively. Each combinatorial library was subjected to four rounds of selection under increasingly stringent conditions according to Figure S3-1. Twelve clones from the final mouse library and twelve clones from the final human combinatorial library were sequenced by conventional Sanger methods, with the mouse library yielding three unique sequences and the human library yielding six unique sequences (Table 3-1). Curiously, several highly enriched mutations including A9F and A9M in the human combinatorial library were completely unrepresented in the final third-generation sequences, replaced instead by amino acid residues that were only included in the degenerate codon to allow access to the residues that were enriched

Yeast Surface Heterodimer Titrations

Optimized, third-generation mouse binders were displayed on yeast cell surface for titration against b-mIL-2Rβγ and exhibited 100-fold improved binding affinities over IL-2 or Rbg40.1F (Figure 3-9). On the other hand, third-generation human binders exhibited binding affinities only comparable to those of Rbg40.1F (Figure 3-10).

Methods

Creation of SSM libraries

Forward primers and reverse primers were designed for each amino acid residue on Rbg40.1F (Table S3-1), resulting in a "left" PCR product with a degenerate NNK codon and a "right" PCR product when amplified with COR and COF primers, respectively. Amplification of "left" and "right" products by COF and COR primers yielded a series of template products each consisting of a degenerate NNK codon at a different residue position. These products were pooled to yield the SSM library. SSM libraries were transformed by electroporation into conditioned *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector, using the protocol previously described by Benatuil *et al*.[57]

Creation of combinatorial libraries

Two separate combinatorial libraries for Rbg40.1F were created by assembly PCR, one specific for mIL-2Rβ/γ_c and one specific for hIL-2Rβ/γ_c. Oligonucleotides used are shown in Table S3-2. Genes were transformed by electroporation into conditioned *Saccharomyces cerevisiae* strain EBY100 cells, along with linearized pETCON vector, using the protocol previously described by Benatuil *et al*.[57]

Fluorescence-activated cell sorting

Fluorescence-activated cell sorting was conducted using mouse or human IL-2Rβ and γ_c receptors, expressed as acid/base zipper heterodimers with C-terminal BAP tag (Table S3-3 and Table S1-4, respectively) to facilitate biotinylation (b-mIL-2Rβγ or b-hIL-2Rβγ), provided as gifts from the Garcia Lab. Yeast cells containing SSM or combinatorial libraries were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm and induced in 1 mL SGCAA for expression at 30°C for 16 h. Five million (5×10^6) cells were incubated with b-mIL-2Rβγ or b-hIL-2Rβγ for 2 h in phosphate buffered saline supplemented with 0.5% BSA and 2 mM EDTA (PBE); and labeled with 5 μg streptavidin, Alexa Fluor 647 conjugate (SA-647) and 5 μg FITC conjugated chicken anti-C-Myc (anti-C-Myc-FITC, Life Technologies) in 250 μL PBE for 15 min on ice; and collected in 1 mL c-Trp-Ura media based on a combination of gates based on FSC, SSC, PE, and APC signals. Selection process was repeated for four rounds with concentration of b-mIL-2Rβγ or b-hIL-2Rβγ and incubation temperature varying as indicated in Figure S3-1. For rounds of selections involving a dissociation step, cells were incubated with 1 mL PBE for 1 h at 37°C after incubation with b-mIL-2Rβγ or b-hIL-2Rβγ but before labeling with SA-647 or anti-C-Myc-FITC. FACS experiments were performed on the SH800 Cell Sorter from Sony Biotechnology equipped with a 488 nm excitation laser.

Yeast-display titrations

Yeast cells were grown in 1 mL c-Trp-Ura media at 30°C for 24 h or to an optical density of 1.0 at 600 nm, and induced in 1 mL SGCAA for expression at 30°C for 16 h. Fifty thousand (5×10^5) cells were collected, incubated with b-hIL-2Rβγ for 2 h at 37°C in PBE, and labeled with 1 μg SA-647 and 1 μg anti-C-Myc-FITC in 50 μL PBE for 15 min on ice. Human IL-2 displayed on EBY100 strain cells containing pCT302 vector[45] and Rbg40.1F displayed on

EBY100 strain cells containing pETCON vector served as controls where applicable. Measurements were performed on Accuri C6 flow cytometer. Cells were gated using FlowJo v10 software and median FL-4 values were used for titration data. Data were fitted using Prism 6 software, using a non-linear one site total saturation binding model.

Library preparation and sequencing

DNA from SSM libraries were prepped for sequencing using a protocol adapted from Chevalier *et al*. Briefly, aliquots of culture from naïve and selected libraries were grown in 1 mL c-Trp-Ura media at 37°C overnight or to an optical density of 4.0 at 600 nm. Cells were treated with Zymolyase to lyse their cell walls (Zymo research), subjected to Miniprep to isolate plasmid DNA (Qiagen), and purified of residual genomic or ssDNA with Exonuclease I and Lambda Exonuclease (New England Biolabs). Two PCR steps were then included to add Illumina adapter regions and to introduce unique barcodes for each library. Sequencing was performed using the MiSeq desktop sequencer.

Sequence-Function Maps

For naïve and selected libraries, forward and reverse paired-end FASTQ files from MiSeq were merged with PEAR[61] using default p-value, minimum overlap, assembly length, and quality score threshold options. For each sequence in the merged FASTQ file, the sub-sequence between restriction sites CATATG and CTCGAG was identified as the gene, checked for appropriate length compared to the SSM template, and translated into the corresponding amino acid sequence. The number of occurrences of each amino acid sequence was tabulated, with only wild-type and single-point mutations included in the final analysis. Enrichment values were determined as follows:

$$E_{s,N} = \log_2(\frac{p_{s,N}/p_{s,0}}{p_{wt,N}/p_{wt,0}})$$

$$p_{i=s,j} = \frac{c_{s,j} + k}{\sum_i c_{i,j} + k}$$

where the enrichment value $E_{s,N}$ of a particular sequence in round N was given by the binary logarithm of the ratio of $p_{s,N}$, the proportion of sequence in round N, and $p_{s,0}$, the proportion of sequence in the naïve library, normalized by the corresponding ratio for the proportion of wild-type template sequence $p_{wt,N}$ in round N over its proportion in the naïve library. Proportion values $p_{s,j}$ were calculated by taking the ratio between the number of observations of sequence s in round j, plus a pseudocount factor k = 0.1, over the number of observations of all wild-type and single mutants, plus a pseudocount factor k = 0.1.

Recombinant expression

Plasmids containing genes in pETCON yeast expression vector were extracted from yeast EBY100 cells using Zymolyase from Zymo Research. Genes of interest were amplified from pETCON with primers shown in Table S3-4, inserted into pET29b *E. coli* expression vector using Gibson assembly[58], and transformed into *E. coli* XL10-gold strain cells by heat shock. Plasmids containing genes encoding designs of interest in pET29b vector were extracted from *E. coli* XL10-gold cells by Miniprep (Qiagen) and transformed into *E. coli* BL21(DE3) strain cells. Cells were grown at 37°C in 0.5 L Terrific Broth to an optical density of 0.6 at 600 nm and induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for expression at 18°C for 6 h. Cell pellets were harvested by centrifugation and lysed by sonication. Protein was purified by Ni-NTA chromatography (eluted in PBS with 250 mM imidazole), verified by SDS-PAGE and

mass spectrometry, and again purified for monomers via size exclusion chromatography (eluted in PBS).

Discussion

In retrospect, several inconsistencies were noted in the sequence-function maps generated and the sequences that emerged from selection against the combinatorial libraries for the third-generation four-helix binders. Significantly improved mIL-2Rβ/γ_c binders have been isolated despite these problems, but the optimization of better hIL-2Rβ/γ_c binders appears to have made no significant progress, suggesting that any issues common to both libraries may have actually hindered the maturation of even stronger mIL-2Rβ/γ_c binders.

The sorting experiments performed on the SSM and combinatorial libraries in this section differ from those in previous sections by the use of streptavidin, Alexa Fluor 647 conjugate rather than streptavidin, phycoerythrin conjugate. The large phycoerythrin fluorophore is prone to non-specific binding, which would incorrectly yield abnormally high affinity (low K_d) values in yeast display titrations. The smaller Alexa Fluor 647 fluorophore is preferable biochemically as it is less susceptible to these issues and would be more appropriate for titrations or selections performed at extremely low concentrations of IL-2Rβ/γ_c target. Unfortunately, the selections were performed with a cell sorter equipped with only a 488 nm excitation laser, which is spectrally incompatible with the excitation spectrum of Alexa Fluor 647. It is believed that the excitation which produced the binding signal in these experiments came not from a monochromatic laser, but rather, accidentally from a FRET-based mechanism: the emission spectrum of the FITC fluorophore used for expression signal overlaps with the absorption spectrum of Alexa Fluor 647. Indeed, no binding signal is observed under the same experimental setup when two known binders are labeled with SA-647 but not anti-C-Myc-FITC.

The advantage of such a setup is that binding signal is indeed coupled to expression within the selected population. However, because the efficiency of FRET transfer is strongly dependent on the distance between the donor and the acceptor and on the relative orientation of the dipoles, the binding signal observed on yeast display is subject to significant statistical variation. As such, enrichment data from the sequence-function maps should be interpreted with appropriate caution for all but the strongest correlations, especially since aberrational depletion of a residue in early rounds of SSM selection result in a bottleneck effect that obscures the true correlation in later rounds.

Tables and Figures

```
mIL2           1 GPGS--------HLEQLLMDLQELLSRMENYRNLKLPRMLTFKFYLPKQA     42
                 .|.|         .||.||:|||.:|:.:.||:|.||.|||||||:||:|
hIL2           1 APTSSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKA     50

mIL2          43 TELKDLQCLEDELGPLRHVLDLTQSKSFQLEDAENFISNIRVTVVKLKGS     92
                 ||||.|||||:||.||..||:|.|||:|.|. ..:.||||.|.|::||||
hIL2          51 TELKHLQCLEEELKPLEEVLNLAQSKNFHLR-PRDLISNINVIVLELKGS     99

mIL2          93 DNTFECQFDDESATVVDFLRRWIAFCQSIISTSPQAAA    130
                 :.||.|::.||:||:|:||.|||.||||||||..
hIL2         100 ETTFMCEYADETATIVEFLNRWITFCQSIISTLT----    133
-----------------------------------------------------------------
mIL2Rβ         1 AVKNCSHLECFYNSRANVSCMWSHEEALNVTTCHVHAKSNLRHWNKTCEL     50
                 ||:..|...|||||||.:||:||.:.||..|:|.|||..:.|.|.|:||||
hIL2Rβ         1 AVQGTSQFTCFYNSRAQISCVWSQDGALQDTSCQVHAWPDRRRWQQTCEL     50

mIL2Rβ        51 TLVRQASWACNLILGSFPESQSLTSVDLLDINVVCWEEKGWRRVKTCDFH    100
                 ..|.||||||||||: |:||.||:||::.:.|:|.|...||.:...||.
hIL2Rβ        51 LPVSQASWACNLILGA-PDSQKLTTVDIVTLRVLCREGVRWRVMAIQDFK     99

mIL2Rβ       101 PFDNLRLVAPHSLQVLHIDTQRCNISWKVSQVSHYIEPYLEFEARRRLLG    150
                 ||:||||:||.||||:|::|.|||||::||.|||.|.:||||||....|
hIL2Rβ       100 PFENLRLMAPISLQVVHVETHRCNISWEISQASHYFERHLEFEARTLSPG    149

mIL2Rβ       151 HSWEDASVLSLKQRQQWLFLEMLIPSTSYEVQVRVKAQRNNTGTWSPWSQ    200
                 |:||:|.:|:|||:|:|:.||.|.|.|.||.|||||..:....|||||||
hIL2Rβ       150 HTWEEAPLLTLKQKQEWICLETLTPDTQYEFQVRVKPLQGEFTTWSPWSQ    199

mIL2Rβ       201 PLTFRTRPADPMKEI    215
                 ||.|||:||...|:.
hIL2Rβ       200 PLAFRTKPAALGKDT    214
-----------------------------------------------------------------
mIL2Rγ         1 PLPEVQCFVFNIEYMNCTWNSSSEPQATNLTLHYRYKVSDNNTFQECSHY     50
                 ||||||||||:|||||||.|||||.|||||||.||.|||:..|:||||
hIL2Rγ         1 PLPEVQCFVFNVEYMNCTWQSSSEPQPTNLTLHYWYKNSDNDKVQKCSHY     50
```

```
mIL2Rγ          51 LFSKEITSGCQIQKEDIQLYQTFVVQLQDPQKPQRRAVQKLNLQNLVIPR    100
                   |||:|||||||:||::|.|||||||||||||::|:|:|.|.|.|||||||.
hIL2Rγ          51 LFSEEITSGCQLQKKEIHLYQTFVVQLQDPREPRRQATQMLKLQNLVIPW    100

mIL2Rγ         101 APENLTLSNLSESQLELRWKSRHIKERCLQYLVQYRSNRDRSWTELIVNH    150
                   |||||||..||||||||.|.:|.: ..||::|||||::.|.||||..|::
hIL2Rγ         101 APENLTLHKLSESQLELNWNNRFL-NHCLEHLVQYRTDWDHSWTEQSVDY    149

mIL2Rγ         151 EPRFSLPSVDELKRYTFRVRSRYNPICGSSQQWSKWSQPVHWGSHTVEEN    200
                   ..:|||||||..||||||||||:||:|||:|.||:||.|:||||:|.:||
hIL2Rγ         150 RHKFSLPSVDGQKRYTFRVRSRFNPLCGSAQHWSEWSHPIHWGSNTSKEN    199
```

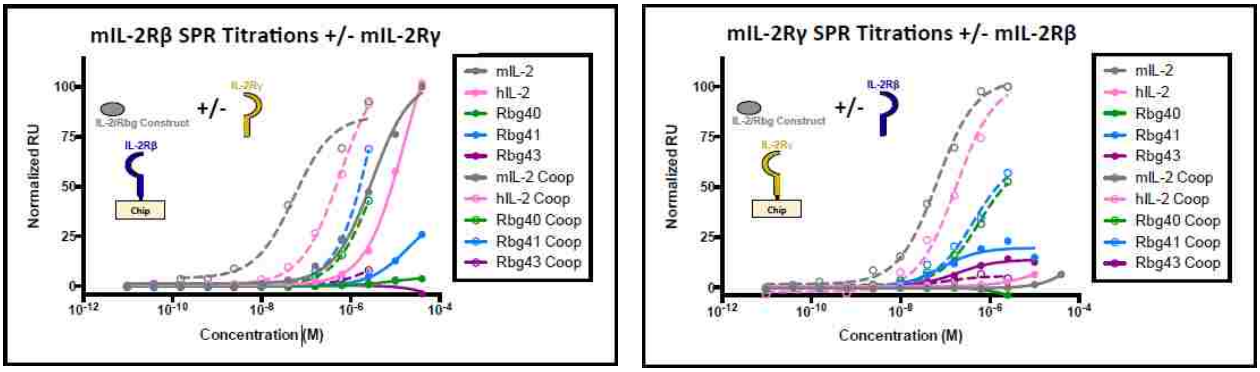
Table 3-1: EMBOSS Pairwise Sequence Alignment[63] results for comparisons of mouse and human IL-2 cytokine, IL-2Rβ ectodomain, and γ_c ectodomain. Sequence similarities were 70.3%, 72.6%, and 85.0% for IL-2, IL-2Rβ, and γ_c, respectively.

Design	Sequence
Rbg40.1F	STKKTQLLAEHALLDALMMLNLLPEPNEKLNRIITTMQSWIFTGKIDGDGAQELAKEVEELEQEHEK RGIDVEDYASNLKVILLELA
Rbg40.2A	STKKTQLHAEQALLDALMMLNLLPEPNEKLNRIITTMQSWIFTGKIDGDGAQELAKEVEELEQEHEK RGIDVENYASNLKVILLELA
Rbg40.2B	STKKTQLHAEQALLDALMMLNLLPEPNEKLNRIITTMQSWIFTGKIDGDGAQELAKEVEELEQEHEK RGIDVENFASNLKVILLELA

Table 3-2: Rbg40.2A and Rbg40.2B differed from Rbg40.1F at 3-4 positions (red) and were designed based on sequence comparisons of mIL-2 and hIL-2.

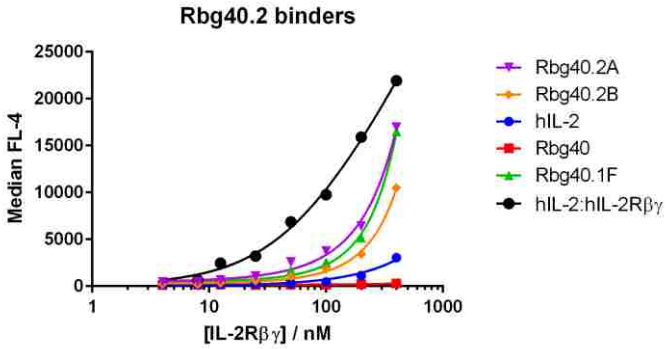
Design	Sequence
Rbg40F.M1	STEKTQLAAEHALRDALMLKHLLNEPNEKLARIITTMQSWQFTGKIDGDGAQELAKEVEELQQEHEV RGIDVEDYASNLKVILLHLA
Rbg40F.M2	STKNTQLAAEDALLDALMLRNLLNEPNEKLARIITTMQSWQFTEKIDGDGAQELAKEVEELQQEHEE RGIDVEDYASNLKVILLQLA
Rbg40F.M3	STEKTQHAAEDALRDALMLRNLLNEPNEKLARIITTMQSWQFTEKIDGDGAQELAKEVEELQQEHEV RGIDVEDYASNLKVILLQLA
Rbg40F.H1	STKKTQLLIEHALLDALDMSRNLPEPNEKLSRIITTMQSWIFTGKIDGDGAQQLAKEVEELEQEHEK RGEDVEDEASNLKVILLELA
Rbg40F.H2	STKKTQLLLEHALLDALHMRRNLPEPNEKLSRIITTMQSWIFTGKIDGDGAQELAKEVEELEQEHEK RGRDVEDDASNLKVILLELA
Rbg40F.H3	STKKTQLLIEHALLDALNMRKKLPEPNEKLSRIITDMQSWIFTGKIDGDGAQQLAKEVEELEQEHEK RGGDVEDYASNLKVILLELA
Rbg40F.H4	STKKTQLLLEHALLDALHMSRELPEPNEKLNRIITDMQSWIFTGKIDGDGAQDLAKEVEELEQEHEK RGGDVEDYASNLKVILLELA
Rbg40F.H5	STKKTQLLIEHALLDALHMSRKLPEPNEKLSRIITTMQSWIFTGKIDGDGAQHLAKEVEELEQEHEK RGGEVEDEASNLKVILLELA
Rbg40F.H6	STKKTQLLIEHALLDALHMKRKLPEPNEKLNRIITNMQSWIFTEKIDGDGAQDLAKEVEELEQEHEK RGQDVEDYASNLKVILLELA

Table 3-3: Third-generation four-helix idealized binders. Rbg40F.M1 to Rbg40F.M3 were evolved against mouse IL-2Rβ/γ_c; Rbg40F.H1 to Rbg40F.H6 were evolved against human IL-2Rβ/γ_c.



Analyte	mIL-2Rβ	mIL-2Rβ (cooperative)	mIL-2Rγ	mIL-2Rγ (cooperative)
mIL-2	3000	53	>40000	66
hIL-2	16000	580	>40000	180
Rbg40	No Binding	3400	No Binding	570
Rbg41	>40000	4500	>40000	420
Rbg43	No Binding	>40000	>40000	>40000

Figure 3-1: Surface plasmon resonance cross-reactivity studies. *Top left*—mIL-2, hIL-2, Rbg40, Rbg41, and Rbg43 were flowed over immobilized mIL-2Rβ in the absence or presence (cooperative) of mIL-2Rγ. *Top right*—mIL-2, hIL-2, Rbg40, Rbg41, and Rbg43 were flowed over immobilized mIL-2Rγ in the absence or presence (cooperative) of mIL-2Rβ. *Bottom*—measured binding affinities for each analyte in nanomolars. Ternary mouse complex affinity for first-generation four-helix mouse mimetics is reduced compared to hIL-2 and significantly reduced compared to mIL-2.[51]



Design	K_d / nM
Rbg40.2A	22694
Rbg40.2B	17610
hIL-2	>100000
Rbg40	No binding
Rbg40.1F	16950
hIL-2:hIL-2Rβγ (positive)	161.7

Figure 3-2: Yeast surface titration curves for Rbg40.2 and Rbg40.2B against mIL-2Rβ/γ_c. Sequence alignment-based substitutions did not improve the affinity of these binders beyond that of unsubstituted second-generation Rbg40.1F, but Rbg40.1F does appear to bind mIL-2Rβ/γ_c more strongly than hIL-2 or first-generation Rbg40. Titration curve for hIL-2 against b-hIL-2Rβγ (hIL-2:hIL-2Rβγ) is presented as a positive control for binding signal.

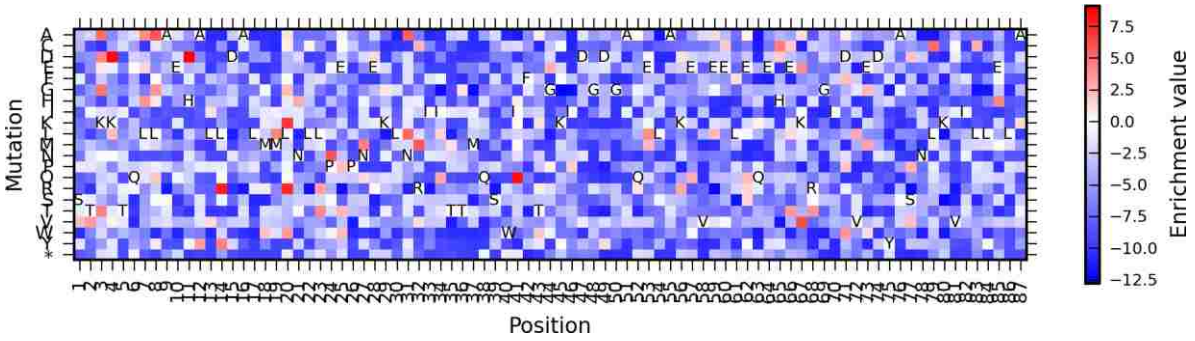


Figure 3-3: Enrichment data for Rbg40.1F SSM round 4 library against mIL-2Rβ/γ_c. The selection for the heatmap above was performed binding against 500 pM b-mIL-2Rβγ at 37°C. Enrichment data for rounds 1-3 and for other temperatures not presented.

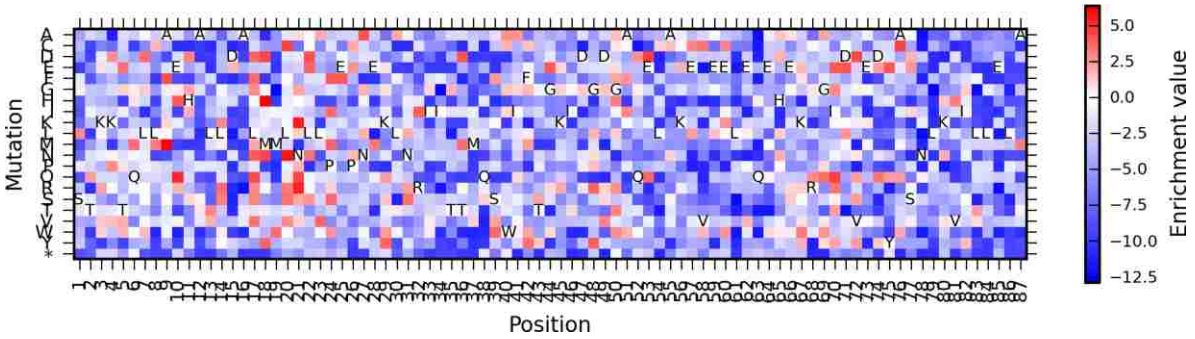


Figure 3-4: Enrichment data for Rbg40.1F SSM round 4 library against hIL-2Rβ/γ_c. The selection for the heatmap above was performed binding against 200 pM b-hIL-2Rβγ at 37°C. Enrichment data for rounds 1-3 and for other temperatures not presented.

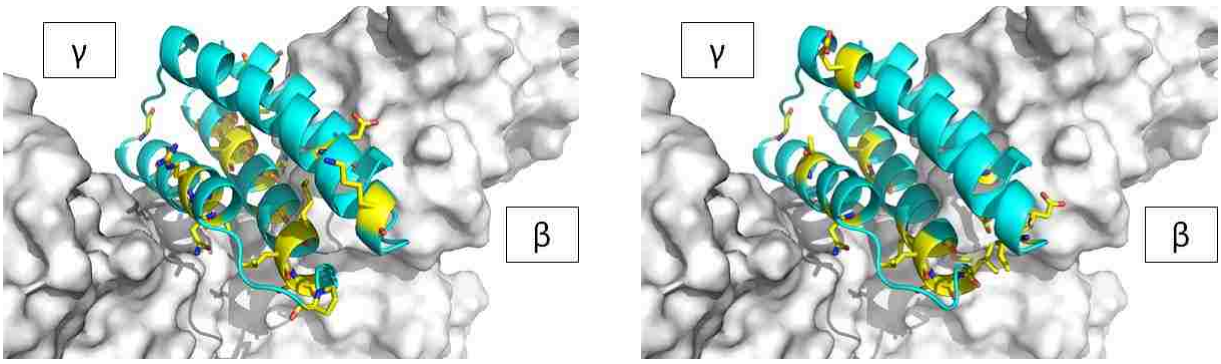


Figure 3-5: Enriched mutations from Rbg40.1F SSM library were designed into combinatorial libraries against mouse IL-2Rβ/γ_c (top left) or human IL-2Rβ/γ_c (top right). Residue positions containing enriched mutations included in the combinatorial library are depicted in yellow. Note the distribution of such positions is nearer the IL-2Rβ or γ_c interfaces in the library selected against mouse IL-2Rβ/γ_c. Computational models visualized using PyMOL.[60]

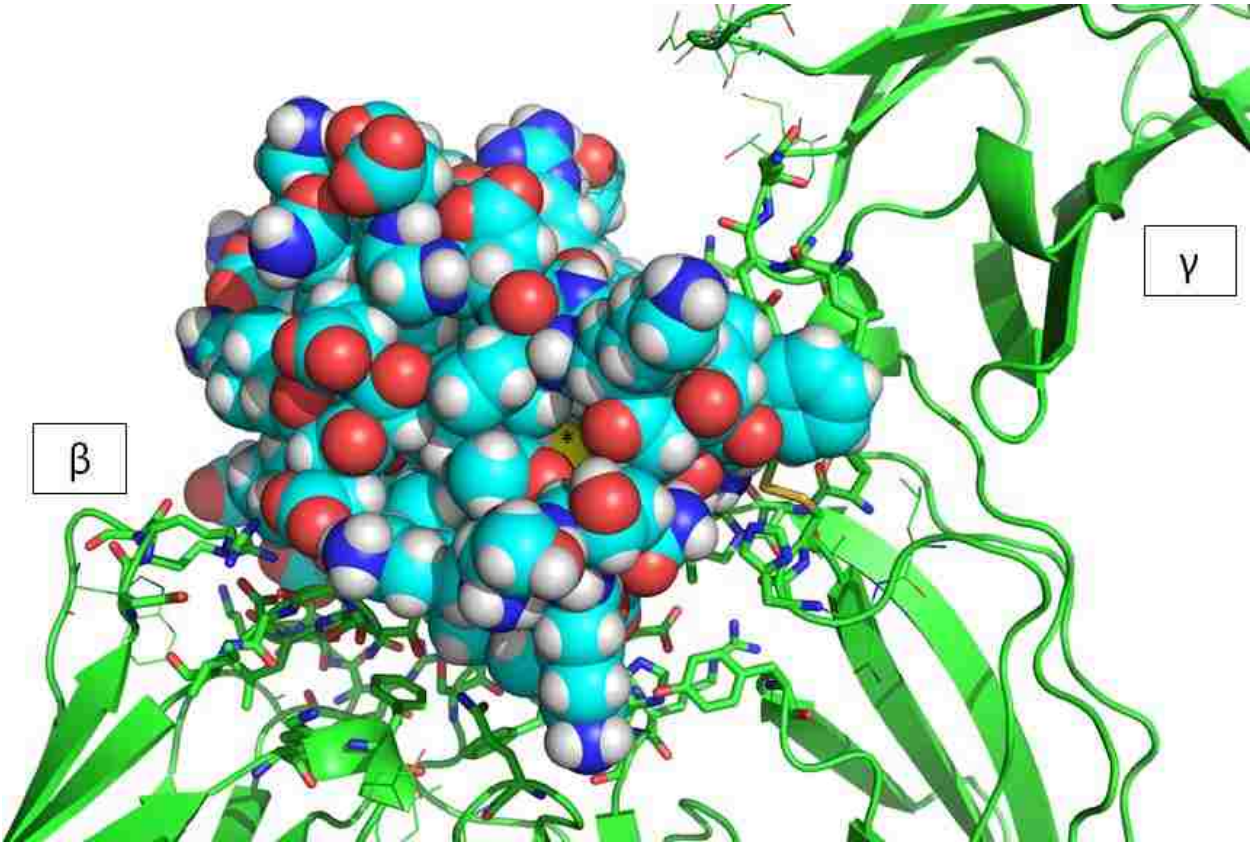


Figure 3-6: Highly enriched mutations on Rbg40.1F SSM library selection against human IL-2Rβ/γ_c include A9M and A9F. Replacing Ala-9 (yellow sphere with black asterisk) with a larger residue allows for better hydrophobic packing and elimination of the void in the protein core. Computational model visualized using PyMOL.[60]

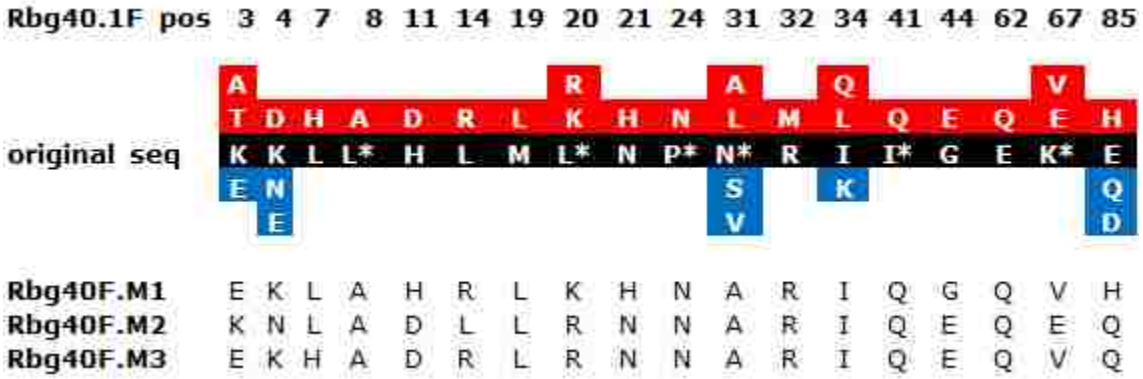


Figure 3-7: Third-generation four-helix binders against mouse IL-2Rβ/γ_c. Combinatorial library for Rbg40.1F contained 18 mutations from original Rbg40.1F with a combined amino acid diversity of 1.05×10^6. Shown in black background are the amino acid identities of the original sequence. Red background indicates desirable amino acids allowed by the designed codon; blue background indicates unintended amino acids allowed by the designed codon. Amino acids from the original sequence are allowed by the designed codon unless indicated by an asterisk. Identities of the three clones (Rbg40F.M1 to Rbg40F.M3) at indicated positions are listed at the bottom.

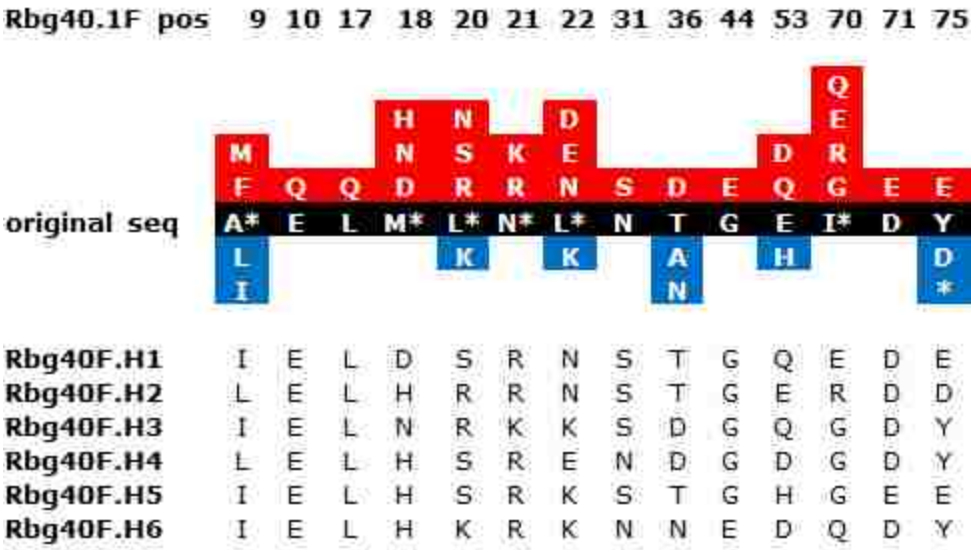
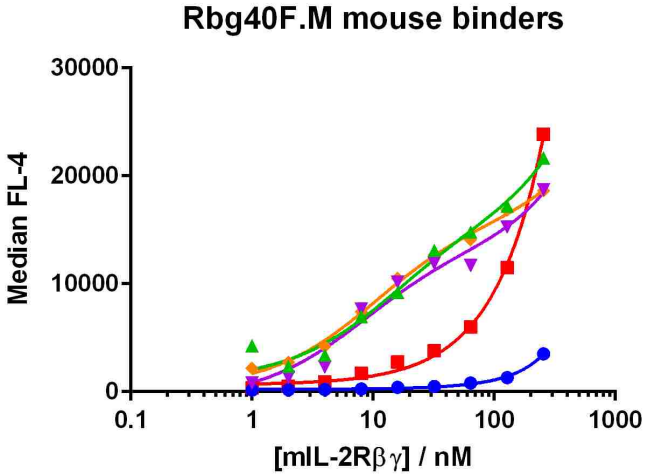
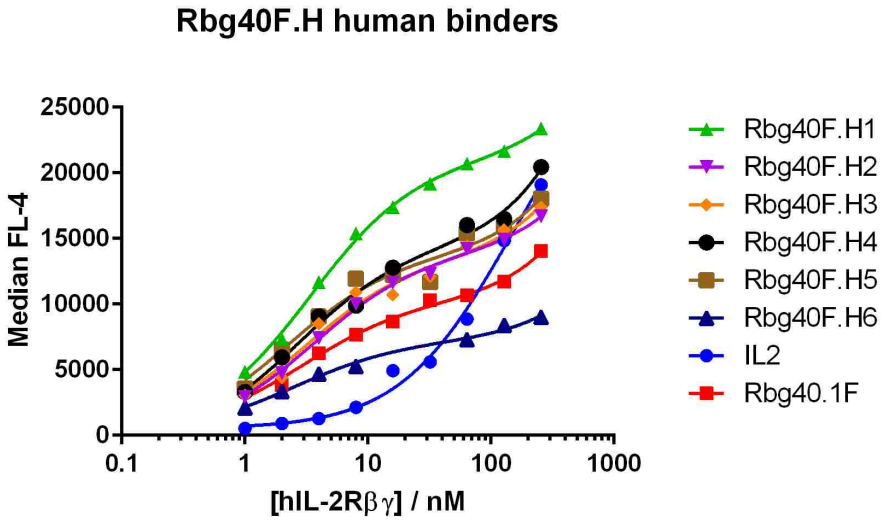


Figure 3-8: Third-generation four-helix binders against human IL-2Rβ/γ_c. Combinatorial library for Rbg40.1F contained 14 mutations from original Rbg40.1F with a combined amino acid diversity of 3.15×10^6. Shown in black background are the amino acid identities of the original sequence. Red background indicates desirable amino acids allowed by the designed codon; blue background indicates unintended amino acids allowed by the designed codon. Amino acids from the original sequence are allowed by the designed codon unless indicated by an asterisk. Identities of the six clones (Rbg40F.H1 to Rbg40F.H6) at indicated positions are listed at the bottom.



Design	K_d / nM
Rbg40F.M1	13.99
Rbg40F.M2	8.077
Rbg40F.M3	10.05
IL2	>10000
Rbg40.1F	>10000

Figure 3-9: Yeast surface titration curves for third-generation binders against mIL-2Rβ/γ_c with IL-2 and Rbg40.1F control. All three clones expressed successfully with dramatically improved affinity for mIL-2Rβγ heterodimer compared to IL2 or Rbg40.1F.



Design	K_d / nM
Rbg40F.H1	3.373
Rbg40F.H2	3.624
Rbg40F.H3	3.056
Rbg40F.H4	2.64
Rbg40F.H5	2.215
Rbg40F.H6	2.484
IL-2	120.6
Rbg40.1F	2.89

Figure 3-10: Yeast surface titration curves for third-generation binders clones against hIL-2Rβ/γ_c with IL-2 and Rbg40.1F control.

Supplementary Materials

Computational Design of Interleukin-2 Mimetics

1a15A	1a19A	1a4pA	1a70A	1a7gE	1a8oA	1a9nB	1aazA	1ab1A	1abaA	1ahoA	1ailA
1akhA	1avsA	1awpA	1ay7B	1b2sD	1b2uD	1b33N	1b3aA	1b3sD	1b4bA	1b4fA	1b8iA
1b8iB	1bc8C	1bcgA	1bcpF	1beoA	1bf4A	1bhpA	1bjaA	1bm8A	1boxA	1brsD	1bt0A
1bwoA	1bxiA	1bxyA	1c48A	1c4qA	1c6rA	1c75A	1cc5A	1cc8A	1ccdA	1cf7A	1cf7B
1cgjI	1chzA	1ci4A	1cktA	1cm3A	1cmiA	1cnoA	1cnrA	1crnA	1cseI	1ctfA	1ctjA
1cxyA	1cxzB	1cy5A	1cyjA	1cyoA	1d1iA	1d1kA	1deeG	1dfuP	1di2A	1dj8A	1djtA
1dk1A	1dokA	1dp7P	1dsxA	1dszA	1dszB	1dt4A	1dtjA	1dtxA	1du0A	1duxC	1dxsA
1e0bA	1eayC	1ec6A	1em7A	1enhA	1eodA	1eoeA	1eofA	1eqtA	1es1A	1etkA	1etoA
1eueA	1euvB	1f1fA	1f60B	1f80D	1f9fA	1f9pA	1f9qA	1f9rA	1f9sA	1fiaA	1fipA
1fjlA	1fk5A	1fm0D	1fr2A	1fs1A	1fseA	1fxdA	1g1xB	1g2rA	1g8qA	1g9oA	1gdvA
1gjsA	1gq5A	1guuA	1gv5A	1gvdA	1gvnA	1gxtA	1gxuA	1h3lA	1h75A	1hb6A	1hbkA
1hstA	1hypA	1hz6A	1i27A	1i2tA	1i70A	1i8vA	1iccA	1ig5A	1ig7A	1ihjA	1iqzA
1iv9A	1j1vA	1j34C	1j75A	1je8A	1jg5A	1jggA	1ji7A	1jk4A	1jkoC	1jo0A	1jyrA
1k50A	1k51A	1k52A	1k53A	1k61A	1k8uA	1k96A	1ka8A	1kh0A	1kkmH	1kp6A	1kq1A
1ku3A	1kuqA	1kv0A	1kw4A	1kwaA	1l0iA	1l9aA	1l9lA	1latA	1lddA	1le8A	1le8B
1lfbA	1lfdA	1lj0A	1lkyA	1lkyB	1lliA	1lmb3	1ln0A	1lngA	1lniA	1lp1A	1lp1B
1lqxA	1lr6A	1lriA	1ls9A	1lw6I	1m20A	1m2iA	1mgrA	1mhhE	1mhxA	1mi0A	1midA
1mj4A	1mk0A	1mn8A	1mo1A	1n0qA	1n7eA	1na3A	1nfjA	1nh9A	1nr4A	1nu4A	1nviD
1o82A	1oaiA	1oeyA	1ogwA	1ohzB	1oiaA	1okhA	1omyA	1opdA	1orcA	1osdA	1p7iA
1p7jA	1pchA	1pdrA	1pgbA	1pgxA	1pk1A	1pk1B	1pk3A	1pohA	1ptfA	1pueE	1pufB
1pylA	1pytA	1q5yA	1q8hA	1qavA	1qb5D	1qbjA	1qdvA	1qe6A	1qx2A	1qzmA	1r1qA
1r4pB	1r69A	1r6jA	1r7jA	1r8hA	1rioH	1rzlA	1rzxA	1s12A	1s29A	1scjB	1segA
1sfuA	1sibI	1sj1A	1sknP	1snbA	1sphA	1stfI	1sv0A	1sv0C	1t07A	1t1vA	1t2hA
1t2iA	1t6oA	1t7aA	1t7bA	1t7eA	1t8kA	1tbxA	1tc3C	1tgrA	1tigA	1tm1I	1tm4I
1tm5I	1tm7I	1tmgI	1to1I	1to2I	1ttzA	1tukA	1u0sA	1u84A	1u9lA	1u9mA	1u9uA
1ucrA	1uj8A	1ujzA	1ulrA	1unkA	1usmA	1usoA	1utgA	1v5iB	1v74B	1vbwA	1vjkA
1vjqA	1vjwA	1vkuA	1vmgA	1vqoS	1vqoX	1vzmA	1w0tA	1w2iA	1w2lA	1w53A	1w85I
1whzA	1wm3A	1wmhA	1wmhB	1wmiA	1wrkA	1wudA	1wv9A	1wveC	1wvnA	1x1wD	1x1xD
1x1yD	1x2iA	1x3oA	1x6iA	1x7vA	1xf5L	1xl3C	1xmkA	1xxaA	1y33I	1y34I	1y39A
1y3bI	1y3cI	1y3dI	1y3fI	1y48I	1y4aI	1y4yA	1y51A	1y6xA	1y7yA	1yd0A	1yd2A
1yd3A	1yd4A	1yd5A	1yd6A	1ylqA	1ymhE	1yn3A	1yn4A	1ynrA	1yo5C	1yu5X	1yu7X
1yu8X	1z3eB	1z96A	1zeiA	1zpvA	1zpwX	1ztgA	1zu3A	1zutA	1zveA	1zvgA	1zyvA
1zywA	1zzkA	2a1jA	2a1jB	2a26A	2a6cA	2a6sA	2a7tA	2acyA	2ahmA	2aibA	2algA
2ascA	2atbA	2avpA	2axyA	2b4jC	2b5aA	2b7cB	2b8iA	2b97A	2b9dA	2bayA	2bf3B
2bh1X	2bkfA	2bkyA	2bkyX	2blfB	2bopA	2bosA	2bpsA	2bwbA	2bwfA	2c5rA	2cb8A
2cc6A	2cclB	2ccqA	2cg5B	2ci2I	2cjjA	2ckxA	2cm0A	2cmpA	2co5A	2croA	2cs7A
2cviA	2cwyA	2cx6A	2cz8A	2czsA	2d0sA	2d1pC	2d3dA	2d68A	2d8dA	2djwA	2ds2B
2duyA	2dyjA	2dznB	2eaqA	2ef8A	2efvA	2egdA	2egoA	2ehsA	2ek0A	2ek1A	2epiA
2eq7C	2eq8C	2eq9C	2erhA	2erlA	2ewhA	2ewtA	2exvA	2f0aA	2f3cI	2f3nA	2f4mB
2f5yA	2f60K	2fazA	2fb0A	2fe5A	2fepS	2ffgA	2ffmA	2fi0A	2fiuA	2fklA	2fmaA
2fs1A	2ftxA	2ftxB	2fu4A	2fz6A	2fztA	2g0cA	2gomA	2gsvA	2gtgA	2guzA	2guzB
2gykA	2gzeA	2gzfA	2gzgA	2gziA	2h1kA	2h27A	2hddA	2heoA	2hinA	2hl7A	2hosA
2hpjA	2hprA	2hueC	2hxxA	2hzcA	2i04A	2i08A	2i0xA	2i5uA	2i61A	2i6vA	2i89A
2ibjA	2ictA	2idoB	2igdA	2j5yA	2j9uA	2nn4A	2nojB	2nptB	2ns0A	2nyzD	2nzcA
2nzuL	2o37A	2o3lA	2o4tA	2o4xB	2o6kA	2o6vB	2o6vD	2o8xA	2o9uX	2obpA	2ochA
2ocsA	2odmA	2ofyA	2on8A	2onqA	2oo2A	2oo9A	2ooaA	2ovgA	2ovoA	2oxlA	2oy9A
2ozfA	2p13A	2p5kA	2p5mA	2p7vB	2p9xA	2pa1A	2pf5A	2pg4A	2pktA	2plhA	2pmrA
2pntA	2posA	2ppxA	2q3gA	2q5wD	2q79A	2q9vA	2qarA	2qb1A	2qffA	2qg1A	2qhoB
2qifA	2qjlA	2qklB	2qkqA	2qkvA	2qmtA	2qn6B	2qnwA	2qsbA	2qswA	2qwoB	2qzgA
2r0rA	2r1jL	2r1qA	2r2zA	2ra4A	2rgiA	2rhfA	2rilA	2rjiA	2rjvA	2rk5A	2rknA
2sn3A	2tecI	2ux9A	2uyzB	2uzcA	2v08A	2v1wA	2v3sA	2v6xA	2v6yA	2v6yB	2v8sV

2vdbB 2vh7A 2vi6A 2vjeA 2vjeB 2vlqA 2vn5B 2vn6B 2vqeO 2vqgA 2vx9A 2vxaA
2vyxA 2w2uA 2w4fA 2w7vA 2w9jA 2w9qA 2wiuB 2wptA 2wy8Q 2x1fA 2x2bA 2x3dA
2x48A 2x4kA 2x5gA 2xcjA 2xenA 2xf5A 2xf7A 2xi8A 2xj3A 2xk5B 2xkoC 2xmjA
2xmmA 2xmwA 2xnqA 2xnrA 2xqqA 2xrhA 2xvcA 2xvtA 2xz0D 2y1rI 2y3yA 2yeoA
2yizA 2yj0A 2ywkA 2yy3A 2za4B 2zbcA 2zboA 2zc2A 2zdjA 2zfzA 2zggA 2zonG
2zplA 2zqeA 2zw0A 2zw1A 2zxyA 3a01A 3a02A 3a03A 3a0sA 3a0vA 3a1pB 3a1yA
3a4rA 3a5eA 3a9fA 3a9jA 3a9jB 3adgA 3adjA 3aeiA 3ajfA 3ajiB 3aqeA 3arcU
3aucA 3ax2A 3ax3A 3b0fA 3b1lX 3b1sB 3b4qA 3b7hA 3bbzA 3bd1A 3beeA 3bp3A
3bpdA 3bpqB 3bpuA 3bq7A 3bqpA 3br8A 3briA 3brlA 3bs3A 3bs5B 3bs7A 3bs9A
3bypA 3bzvB 3bzxB 3bzyB 3bzzB 3c00B 3c01E 3c0fB 3c57A 3c8pA 3ccdA 3cedA
3cfiA 3ci0I 3cimA 3cjkB 3cjsB 3cnqP 3cqxC 3csxA 3cu4A 3cuoA 3cztX 3d2wA
3d36C 3d3bJ 3d3cJ 3d5rC 3d5sC 3d8lA 3dbhI 3dblI 3df6A 3dgpA 3dgpB 3dmiA
3dnjA 3dr0A 3ds0A 3ds1A 3ds2A 3ds4A 3dsgA 3dvhA 3dwgC 3dxsX 3e19A 3e21A
3e7lA 3e7rL 3eabA 3eipA 3ej3A 3ej7A 3ej9A 3ej9B 3encA 3eqyA 3eusA 3excX
3eyiA 3f2eA 3f6wA 3fapB 3fauA 3fb9A 3fblA 3ff5A 3filA 3fjuB 3fmyA 3fojA
3fryA 3fyaA 3fyrA 3g0vA 3g1bA 3g21A 3g26A 3g27A 3g28A 3g6eI 3g9mA 3ge3C
3ggeA 3ghdA 3gjnA 3gklC 3gzfA 3gzmA 3h2vE 3h36A 3h8yA 3h92A 3hglA 3hluA
3hmrA 3hphE 3htuA 3hz7A 3i31A 3i71A 3ibwA 3ic4A 3id1A 3id3A 3idwA 3il8A
3ipjA 3iqoA 3iufA 3iwlA 3jstA 3jtnA 3jtpA 3jtzA 3jxoA 3k1rB 3k2aA 3k2tA
3kr3D 3kseD 3ku7A 3kucB 3kudB 3kupA 3kw6A 3kwrA 3kycD 3kz3A 3kzdA 3l7hA
3l9aX 3ldcA 3lddA 3le1A 3lfpA 3lhrA 3lklA 3llhA 3lnyA 3lo3A 3lpeA 3lvkB
3m03A 3m1dA 3m4gA 3m8aA 3ma5A 3mabA 3mb2A 3md1A 3mdfA 3mi7X 3mshA 3mszA
3mtnB 3mxgA 3mycA 3mzwB 3n01A 3n3kB 3n7sC 3narA 3nerA 3ng8A 3nirA 3nnhA
3ns6A 3ntwA 3nxaA 3nzlA 3o1fA 3o2pE 3o46A 3oeiC 3offA 3ofgA 3ofhA 3oizA
3olmD 3omtA 3onjA 3oufA 3ousA 3ouvA 3ov5A 3ov8A 3ozzB 3p04A 3p1xA 3p5tL
3ph0C 3ph2B 3phxB 3pluA 3po0A 3psmA 3pxmA 3pyjA 3q2pA 3qc8B 3qmqA 3qmxA
3qwoC 3qx1A 3r07C 3r15A 3r27A 3r2cJ 3r68A 3r69A 3r8sZ 3rcoA 3rd2A 3rdyA
3rjpA 3rjsA 3rklA 3rkqA 3rq9A 3rzwC 3s1kA 3s64A 3s7rA 3s8qA 3sd6A 3sjmA
3solA 3szsA 3t46A 3t47A 3t49A 3t7yA 3tduC 3tjyA 3tnvA 3toeA 3toqA 3tqmA
3tr3A 3trgA 3tysA 3u43A 3u5ek 3u8vA 3ub0B 3ucgA 3ue7B 3ul4B 3ulqB 3v2cR
3v2d2 3v3lA 3v6cB 3v6eB 3v7eA 3vc8A 3vcbA 3vkeA 3ygsP 3zr8X 3zzpA 451cA
4a20A 4a3nA 4a47A 4a4jA 4a56A 4a5uB 4a5xA 4a69C 4a8xA 4aazA 4ajyC 4b2bB
4b6wA 4b6xA 4dfcA 4dh2B 4dijA 4djgA 4ds1A 4dtgK 4duqA 4e1pA 4e34A 4e5xG
4e6kG 4eefG 4eewA 4efoA 4ep8A 4ersA 4es1A 4evuA 4f3fC 4f87A 4fisA 4fjoA
4fl4A 4fm3A 4fp5D 4fqnA 4fxiA 4g3oA 4g69A 4gd1F 4gd1R 4gd1T 4gdkA 4gocA
4gofA 4gu2A 4gvbB 4heiA 4hjkA 4hk2A 4htiA 4hw5A 4i16A 4i4dA 4id3A 4iejA
6insE 9antA

Table S1-1: List of 1046 scaffolds used in MotifGraft, with first four characters corresponding to PDB ID and fifth character corresponding to chain ID.

Scaffold	Input helices
1mw5A	A, D
1tqgA	A, C
1u89A	A, D
1wrdA	A, C
1x91A	A, D
1yzmA	A, D
1z0kB	A, D
1z6oM	A, D
2j0oA	A, C, D
2lm9A	A, D
2p5tA	A, D

2rldA	A, D
2v0oA	A, D
2v6yA	A, D
3ajfA	A, D
3am6A	A, D
3dfbA	A, D
3eabA	A, D
3iqcA	A, D
3onjA	A, D
3pcvA	A, D
3rf3A	A, D
4hwhA	A, D

Table S1-2: List of scaffolds identified by TM-align[49] as structurally compatible with IL-2, with corresponding IL-2 input helices provided. Four of these entries are redundant with those in Table S1-1.

Oligonucleotide	**Sequence**
lib_Rbg03_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGGAGTACCTGGG
lib_Rbg03_r2	CAGGTTCTGCAGCAGTTCTTTGGTTTCGTCCGCGAAAACACCCAGGTACTCCATATGGCT
lib_Rbg03_f3	AAGAACTGCTGCAGAACCTGAACGACACCCTCCTGGAACTCGAAAAAAACCCGGAAGACA
lib_Rbg03_r4	AGGGTCAGCGCCGCACGGAACGCTTCGTTGATCAGTTCCATGTCTTCCGGGTTTTTTTCG
lib_Rbg03_f5	TTCCGTGCGGCGCTGACCCTGCTGGGTATGGCGGGTACCATGGGTTTCTCTTCTCTCCTG
lib_Rbg03_r6	GTGGTTTTGATTTCAGAGTTAMGGRYAGCGTCCGCAKCAWKCAGCAGGKSTWYGCACWGTWK CAGGAGAGAAGAGAAACCCA
lib_Rbg03_f7	AACTCTGAAATCAAAACCACCTCTGACGCGMTAGACGCGGCTCGTGYGATCGTTGAGGAGAT CACCCGTTTTGTTGATAA
lib_Rbg03_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGAGAAACGATTTTATCAACAAAACGGGTGAT
lib_Rbg12_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGGTTCGTGTGGACCAGA
lib_Rbg12_r2	GACAGTTCGTCGAGCAGTTGCATAACTTCGTTGAACAGGTTCTGGTCCACACGAACCATA
lib_Rbg12_f3	CAACTGCTCGACGAACTGTCCCAAGACATCACCTCTCCGAAAAACGTTCGTAAACTGGCG
lib_Rbg12_r4	GAGATTCGTTTTCCTGAGACAGTTTCGCAGCCGCGTCTTGCGCCAGTTTACGAACGTTTT
lib_Rbg12_f5	GTCTCAGGAAAACGAATCTCTGGACCTGGCGTGCGCGASGGCGATCTCCATGGCTCAGGA
lib_Rbg12_r6	CGGCACGTTCGGGTCAGCGATAGCCTCCTGAGCCATGGAGATCGC
lib_Rbg12_f7	TCGCTGACCCGAACGTGCCGRMAARAGGTCGTWYAGACCTGRWASMCATCCTGTYAGMCCTG GAAGCTATTTCCCTCGA
lib_Rbg16_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCCTCTGAGATGTCT
lib_Rbg16_r2	TCAGGCAGAAAGACGGGTTCAGGGTTTTGTCGCAGATAGTAGACATCTCAGAGGACATAT
lib_Rbg16_f3	GAACCCGTCTTTCTGCCTGAAATTCCTGAACACCAAATTCGCGTCCGCTAACCTCCAGGC
lib_Rbg16_r4	CGCACGCGCCTGAGTAGAATCGAGGGTGGTCTTCGCGAGCGCCTGGAGGTTAGCGGACGC
lib_Rbg16_f5	ATTCTACTCAGGCGCGTGCGAMCCAGGCGATTRCTAAAGCGCAAGCGGCTATCGCTGGTGGT GTTGACCCGGAA
lib_Rbg16_r6	TTACCGATCGCGTTTTGCAGAKCAWSGAGACAAKSCYYGTAAGCCAGTTTAGWTTCCGGGTC AACACCACCAG
lib_Rbg16_f7	CTGCAAAACGCGATCGGTAACGCGGAAGAAGCGTTCGAGCACGCGGCGTCTGGTGACGGC
lib_Rbg16_r8	CACCGTCCAGAGCAGCAGACACTTTCATGTTCGCACCCATGCCGTCACCAGACGCCGCGT
lib_Rbg16_f9	GTCTGCTGCTCTGGACGGTGCGGACTGGTGCCTGGACGCTCTGTCTCGTCTCCGTTCCGT
lib_Rbg16_r10	GTTTTTCAGAGTTTTCGCGTTATTAACAGCAGAAGAGTCAACGGAACGGAGACGAGACAG
lib_Rbg16_f11	ACGCGAAAACTCTGAAAAACCTGTGCGGTATCGCGCTGGTTATCGCGAATATGCTGCCGC
lib_Rbg16_r12	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGGTTACGCGGCAGCATATTCGCGATA

lib_Rbg12_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGGGAAATAGCTTCCAG
lib_Rbg17_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGCACCGTTCTTGCCGTAATTCTATGCGTCAG
lib_Rbg17_r2	CCAGCGCCTGCAGAGACGCACCGATCGCCATCTGGATCTGCTGACGCATAGAATTACGGC
lib_Rbg17_f3	TGCGTCTCTGCAGGCGCTGGCGATGGGCGCTCACGCGTCTAAGGACGTTGTTAACCGTCC
lib_Rbg17_r4	TTCTTCAGACGCCGCGTCAAACGCGAGCTGCGCAACACCCGGACGGTTAACAACGTCCTT
lib_Rbg17_f5	TTGACGCGGCGTCTGAAGAACGTGAACACGCGATGAAGCTGATCGAACTGCTGCTGATGC
lib_Rbg17_r6	ACCTGGAGGAGAGAAGAAACGTCGTTGGTCAGTTCACCACGCATCAGCAGCAGTTCGATC
lib_Rbg17_f7	GTTTCTTCTCTCCTCCAGGTTCGTCCGCCGACCCGTACCTCCTGGAAAGGTGGTGTTGAAGC GCTGGAG
lib_Rbg17_r8	GCGCAGACTCGGTGATCGCCTGTTCSAWASTTRBCGCGYGCTCCAGCGCTTCAACACCAC
lib_Rbg17_f9	GGCGATCACCGAGTCTGCGCGTAATGTTATCAAAGCGTGCGAAGACGACTCTGAATTTAA
lib_Rbg17_r10	AACGATGTCACCGGTCAGATAGTCCGCCAGGTGGTACGCGTTAAATTCAGAGTCGTCTTC
lib_Rbg17_f11	ATCTGACCGGTGACATCGTTGAASAGCWARWASMCGGTCTGCKTGACSTTCAGGGTAAAGCG TCTACCCT
lib_Rbg17_r12	GAACTCACCCAGAGCTTCGTGACGGTCCATCAGTTTTTTGAGGGTAGACGCTTTACCCTG
lib_Rbg17_f13	ACGAAGCTCTGGGTGAGTTCATCTTCGCTAAAAAGCTCCTCGGTATCGACGTTCTCGAGG
lib_Rbg17_r14	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGAACGTCGATACCG
lib_Rbg18_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGGGTTCTATGGAAGCG
lib_Rbg18_r2	TCCATTCGAAAGCCTGTTTGTGCCAAACACGAACACGTTCCGCTTCCATAGAACCCATAT
lib_Rbg18_f3	CAAACAGGCTTTCGAATGGATCTCTATCGCGCTGCGTATCGATGAAGACGCGAAAGCGGG
lib_Rbg18_r4	GATACCTTTTTTGTACCATTCAATCGCTTGTTCCTTCTGACCCGCTTTCGCGTCTTCATC
lib_Rbg18_f5	AATGGTACAAAAAAGGTATCRMAGCGCTGCAGGCGGCGATTKCGGTTATCGTTACCGGTCAA GG
lib_Rbg18_r6	ACGACGCGCACGTTCACACTGTTCACCTTGACCGGTAACGATAAC
lib_Rbg18_f7	AGTGTGAACGTGCGCGTCGTWYGCAGMWAMWACTGRWGSMCAACCTGSTTGACGCGCTGGAC CGTCTGCA
lib_Rbg18_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGTTCGAGGAGCTGCAGACGGTCCAGCGCGTC
lib_Rbg21_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCCGCGCAGGTTATGCTGG
lib_Rbg21_r2	GCTTTAACCGCCGCAATCGCCAGTTCACGCGCCAGGTCCTCCAGCATAACCTGCGCGGAC
lib_Rbg21_f3	GCGATTGCGGCGGTTAAAGCGGACAAAGAAGGTAAAGTTGAAGACGCGGCGACCTACTAC
lib_Rbg21_r4	CGGTACGCGCAACAGATTCCGGGTACAGAACGATGATCTGTYGAASGKCTWSCAGCGCGKST TYGTAGTAGGTCGCCGCGTCTT
lib_Rbg21_f5	GGAATCTGTTGCGCGTACCGCGTACGAACAGATGRYCAMCGAAGCGCAGARAAKAATCKCGW MTCTCGAAAAAGTTCTGCTCGA
lib_Rbg21_r6	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGCAGAACTTTTTCGAG
lib_Rbg25_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGGGT
lib_Rbg25_r2	CCTGTTTGTGCCAAACACGAACACGTTCAGCTTCCATAGAACCCATATGGCTAGCCGACC
lib_Rbg25_f3	TCGTGTTTGGCACAAACAGGCGTTCGAAGCGATCTCTATCGCGCTGCGTATCGACGAAGA
lib_Rbg25_r4	AGCCGCTTCCGCAGCCTGTTCTTTCTGACCCGCTTTTTCGTCTTCGTCGATACGCAGCGC
lib_Rbg25_f5	AACAGGCTGCGGAAGCGGCTRAASMTGGTCTGSWGGMCDYACRAMWGGGTATCGCGGTTATC GTTACCGGT
lib_Rbg25_r6	TCGCCTGCAGACGACGCGCACGTTCGCACTGTTCACCTTGACCGGTAACGATAACCGCGA
lib_Rbg25_f7	TGCGCGTCGTCTGCAGGCGAAAATGATGACCAACVTCRYCATGGCTCAAGCGCGTATCTC
lib_Rbg25_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGTTCCAGCAGAGAGATACGCGCTTGAGCCAT

Table S1-3: Oligonucleotides used in the assembly of focused MotifGraft libraries. All sequences are written in 5'→ 3' orientation.

Human IL-2Rβ with basic leucine zipper

AVNGTSQFTCFYNSRANISCVWSQDGALQDTSCQVHAWPDRRRWNQTCELLPVSQASWAC
NLILGAPDSQKLTTVDIVTLRVLCREGVRWRVMAIQDFKPFENLRLMAPISLQVVHVETH

RCNISWEISQASHYFERHLEFEARTLSPGHTWEEAPLLTLKQKQEWICLETLTPDTQYEF
QVRVKPLQGEFTTWSPWSQPLAFRTKPAALGKDTSRGGLEVLFQGPEFGGSTTAPSAQLK
KKLQALKKKNAQLKWKLQALKKKLAQHHHHHH

Human γ_c with acidic leucine zipper

LNTTILTPNGNEDTTADFFLTTMPTDSLSVSTLPLPEVQCFVFNVEYMNCTWNSSSEPQP
TNLTLHYWYKNSDNDKVQKCSHYLFSEEITSGCQLQKKEIHLYQTFVVQLQDPREPRRQA
TQMLKLQNLVIPWAPENLTLHKLSESQLELNWNNRFLNHCLEHLVQYRTDWDHSWTEQSV
DYRHKFSLPSVDGQKRYTFRVRSRFNPLCGSAQHWSEWSHPIHWGSNTSKENSRGGLEVL
FQGPEFGGSTTAPSAQLEKELQALEKENAQLEWELQALEKELAQGLNDIFEAQKIEWHEH
HHHHH

Table S1-4: b-hIL2Rβγ heterodimer construct was formed by co-expressing IL-2Rβ ectodomain (residues 1-214) and γ_c ectodomain (residues 1-232) onto complementary leucine zipper base/acid pairs, with the latter containing a C-terminal BAP tag to facilitate biotinylation.

Primer	Templates	Sequence
5'_Rbg32	Rbg32, Rbg32.1 to Rbg32.4, Rbg32.6	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCCGCGCAGGT TATG
3'_Rbg32	Rbg32, Rbg32.1 to Rbg32.4, Rbg32.6	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCAGAACTTTTT CCAG
3'_Rbg32_C71	Rbg32.5, Rbg32.7	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCAGAACTTTTT CGC
5'_Rbg35	Rbg35	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAAA ATGGCAAC
3'_Rbg35	Rbg35	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGGCACGCTTC
5'_Rbg36-41	Rbg36 to Rbg41	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAAA ACTCCAGC
3'_Rbg36-39	Rbg36 to Rbg39	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCTTGAGTTCCA G
3'_Rbg40-41	Rbg40, Rbg41	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGCCAGCTCC
5'_Rbg42-44	Rbg42 to Rbg44	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAAA GATCCAGC
3'_Rbg42-44	Rbg42 to Rbg44	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGAGAGCTCCA G
3'_Rbg39	Rbg39	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCTTGAGCTCCA G
3'_Rbg41	Rbg41	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGCCAGTTCG

Table S1-5: Primers used to subclone genes from first-generation designs into pET29b. All sequences are written in 5'→ 3' orientation.

```
relax.static.linuxgccrelease -database /work/shawnyu/rosetta/main/database -
ignore_unrecognized_res -relax:constrain_relax_to_start_coords -
relax:coord_constrain_sidechains -relax:ramp_constraints false -ex1 -ex2 -
use_input_sc -score:weights talaris2013 -no_his_his_pairE -no_optH false -
flip_HNQ -s $1
```

Protocol S1-1: Representative command for preparation of input PDB by relax with coordinate constraints.

```
rosetta_scripts.static.linuxgccrelease -database
/work/shawnyu/rosetta/main/database -overwrite -out:file:renumber_pdb false -
ex1 -ex2 -ignore_zero_occupancy false -out:output -parser:protocol
motifgraft_AD.xml -s $1 -nstruct 20
```

motifgraft_AD.xml

```
<dock_design>
```

```
    <TASKOPERATIONS>
        <InitializeFromCommandline name="init"/>
        <LimitAromaChi2 name="arochi2"/>
        <IncludeCurrent name="inclcur"/>
        <ExtraRotamersGeneric name="exrot" ex1="1" ex2="1"
extrachi_cutoff="1"/>
        <DisallowIfNonnative name="nohis" disallow_aas="GPH"/>
        <OperateOnCertainResidues name="hotspot_onlyrepack">
            <RestrictToRepackingRLT/>
            <ResiduePDBInfoHasLabel property="HOTSPOT"/>
        </OperateOnCertainResidues>
        <OperateOnCertainResidues name="scaffold_onlyrepack">
            <RestrictToRepackingRLT/>
            <ResiduePDBInfoHasLabel property="SCAFFOLD"/>
        </OperateOnCertainResidues>
        <OperateOnCertainResidues name="context_norepack">
            <PreventRepackingRLT/>
            <ResiduePDBInfoHasLabel property="CONTEXT"/>
        </OperateOnCertainResidues>
        <LayerDesign name="core_design" layer="core" core="40.0"/>
        <RestrictChainToRepacking name="chain1_onlyrepack" chain="1"/>
    </TASKOPERATIONS>
    <SCOREFXNS>
        <talaris2013_sfxn weights="talaris2013">
            <Reweight scoretype="res_type_constraint" weight="1.0"/>
        </talaris2013_sfxn>
        <talaris2013_cart weights="talaris2013">
            <Reweight scoretype="cart_bonded" weight="1.0"/>
            <Reweight scoretype="res_type_constraint" weight="1.0"/>
        </talaris2013_cart>
    </SCOREFXNS>
    <FILTERS>
        <ScoreType name="total_score" scorefxn="talaris2013_sfxn"
score_type="total_score" confidence="0" threshold="0"/>
        <Ddg name="ddg" scorefxn="talaris2013_sfxn" jump="1" confidence="0"
repack="1" repeats="1"/>
    </FILTERS>
    <MOVERS>
        <MotifGraft name="motif_grafting" context_structure="./Rbg.pdb"
motif_structure="./motif-AD.pdb" RMSD_tolerance="1.0"
NC_points_RMSD_tolerance="1.0" clash_score_cutoff="5"
clash_test_residue="GLY"  hotspots="5:12,12:15"
combinatory_fragment_size_delta="3:3,6:1"
max_fragment_replacement_size_delta="0:0,0:0"  full_motif_bb_alignment="1"
allow_independent_alignment_per_fragment="0"
graft_only_hotspots_by_replacement="0"
only_allow_if_N_point_match_aa_identity="0"
only_allow_if_C_point_match_aa_identity="0"
revert_graft_to_native_sequence="1" allow_repeat_same_graft_output="0"/>
        <PackRotamersMover name="pack_graft" scorefxn="talaris2013_sfxn"
task_operations="init,scaffold_onlyrepack,hotspot_onlyrepack,context_norepack
,arochi2,inclcur,exrot,nohis"/>
        <PackRotamersMover name="pack_core" scorefxn="talaris2013_sfxn"
task_operations="init,core_design,hotspot_onlyrepack,context_norepack,arochi2
,inclcur,exrot,nohis"/>
        <TaskAwareMinMover name="cart_min" bb="1" chi="1" jump="1"
cartesian="1" scorefxn="talaris2013_cart"
```

```
task_operations="init,chain1_onlyrepack,context_norepack,arochi2,inclcur,exro
t,nohis"/>
        <TaskAwareMinMover name="kine_min" bb="0" chi="1" jump="1"
scorefxn="talaris2013_sfxn"
task_operations="init,chain1_onlyrepack,context_norepack,arochi2,inclcur,exro
t,nohis"/>
        <FastRelax name="fr" scorefxn="talaris2013_sfxn" repeats="3"
task_operations="init,chain1_onlyrepack,context_norepack,arochi2,inclcur,exro
t,nohis"/>
    </MOVERS>
    <APPLY_TO_POSE>
    </APPLY_TO_POSE>
    <PROTOCOLS>
        <Add mover_name="motif_grafting"/>
        <Add mover_name="pack_graft"/>
        <Add mover_name="pack_core"/>
        <Add mover_name="cart_min"/>
        <Add mover_name="kine_min"/>
        <Add filter="total_score"/>
        <Add filter="ddg"/>
    </PROTOCOLS>
</dock_design>
```

Protocol S1-2: Representative command for running MotifGraft, with an example protocol for input motif consisting of helices A and D shown below.

```
/work/dadriano/DEVEL/rosetta_main_git_epigraft_fragments/main/source/bin/rose
tta_scripts.default.linuxgccrelease -database
/work/dadriano/DEVEL/rosetta_main_git_epigraft_fragments/main/database/ -
overwrite -out:file:renumber_pdb false -ignore_zero_occupancy false -
out:output -parser:protocol ./design_uIdealwResfile.xml -holes:dalphaball
/work/dadriano/PROGRAMS/rosetta-helpers/DAlphaBall.icc -s ./input.pdb -
out:path:all results_D2/ -out:suffix _t89 -nstruct 20
```

design_uIdealwResfile.xml

```
<dock_design>
    <SCOREFXNS>
        <SFXN6 weights="talaris2013.wts"/>
        <SFXN6dA weights="talaris2013_downAla.wts"/>
    </SCOREFXNS>
    <FILTERS>
        <SSPrediction name="sspred"
cmd="/work/dadriano/PROGRAMS/psipred/runpsipred_single" use_probability="0"
use_svm="0" threshold="0.80" confidence="1"/>
        <ScoreType name="rama" scorefxn="SFXN6" score_type="rama"
threshold="0.0" confidence="0"/>
        <PackStat name="pack" threshold="0.63" confidence="1"/>
        <Holes name="holes" threshold="1.2" confidence="0"/>
        <ScoreType name="score" scorefxn="SFXN6" score_type="total_score"
threshold="0.0" confidence="0"/>
        <ResidueCount name="nres" confidence="0"/>
        <CalculatorFilter name="score_res" equation="SCORE/NRES" threshold="-
1.7" confidence="1">
        <SCORE name="SCORE" filter_name="score"/>
        <NRES name="NRES" filter_name="nres"/>
    </CalculatorFilter>
    <CompoundStatement name="filt" >
```

```
        <AND filter_name="sspred"/>
        <AND filter_name="rama"/>
        <AND filter_name="score_res"/>
        <AND filter_name="pack"/>
    </CompoundStatement>
</FILTERS>
<TASKOPERATIONS>
    <InitializeFromCommandline name="init"/>
    <IncludeCurrent name="inclcur"/>
    <LimitAromaChi2 name="limitchi2"/>
    <ReadResfile name="resfile" filename="./input.resfile"/>
</TASKOPERATIONS>
<MOVERS>
    <Dssp name="dssp"/>
    <FastDesign name="fdesign" task_operations="init,resfile,limitchi2"
scorefxn="SFXN6dA" allow_design="1" only_design_worst_region="0"
design_by_psipred="0" design_by_frag_qual="0" repeats="2"
clear_designable_residues="0"  max_redesigns="2000"/>
    <FastRelax name="relax"/>
    <ParsedProtocol name="complexDesign" >
        <Add mover_name="fdesign"/>
        <Add mover_name="relax"/>
        <Add mover_name="dssp"/>
    </ParsedProtocol>
    <LoopOver name="fastDesignProtein" mover_name="complexDesign"
filter_name="filt" drift="0" iterations="10"
ms_whenfail="FAIL_DO_NOT_RETRY"/>
</MOVERS>
<APPLY_TO_POSE>
</APPLY_TO_POSE>
<PROTOCOLS>
    <Add mover_name="fastDesignProtein"/>
    <Add filter_name="sspred"/>
    <Add filter_name="pack"/>
    <Add filter_name="score"/>
    <Add filter_name="score_res"/>
    <Add filter_name="holes"/>
    <Add filter_name="rama"/>
</PROTOCOLS>
</dock_design>


<dock_design>
    <SCOREFXNS>
        <SFXN6 weights="talaris2013.wts"/>
    </SCOREFXNS>
    <FILTERS>
    </FILTERS>
    <TASKOPERATIONS>
        <InitializeFromCommandline name="init"/>
        <LimitAromaChi2 name="limitchi2"/>
        <ReadResfile name="resfile" filename="./input.resfile"/>
    </TASKOPERATIONS>
    <MOVERS>
        <PackRotamersMover name="design_u_resfile"
scorefxn="talaris2013_resCons" task_operations="init,resfile,limitchi2"/>
        <FastRelax name="relax"/>
```

```
    </MOVERS>
    <APPLY_TO_POSE>
    </APPLY_TO_POSE>
    <PROTOCOLS>
        <Add mover_name="design_u_resfile"/>
        <Add mover_name="relax"/>
    </PROTOCOLS>
</dock_design>
```

Protocol S1-3: Representative command and protocol for design of an idealized design after loop remodeling.

```
rosetta_scripts.default.linuxgccrelease -database
/work/shawnyu/rosetta/main/database -ignore_unrecognized_res -overwrite -
out:file:renumber_pdb false -ex1 -ex2 -score:weights talaris2013 \
-s $1 -parser:protocol fast_disulfide.xml -parser:script_vars stringency=$2 -
save_top 1000
```

fast_disulfide.xml

```
<dock_design>
    <SCOREFXNS>
    </SCOREFXNS>
    <FILTERS>
        <ResidueCount name="count_cyd" residue_types="CYD" confidence="1"
min_residue_count="2"/>
        <ScoreType name="dslf_fa13" score_type="dslf_fa13" threshold="0"
confidence="0"/>
        <CalculatorFilter name="dslf" equation="t1*t1" threshold="0.000001">
            <VAR name="t1" filter="dslf_fa13"/>
        </CalculatorFilter>
    </FILTERS>
    <TASKOPERATIONS>
        <RestrictToRepacking name="restrict"/>
        <OperateOnCertainResidues name="cyd_only" >
            <PreventRepackingRLT/>
            <ResidueName3Isnt name3="CYD"/>
        </OperateOnCertainResidues>
    </TASKOPERATIONS>
    <MOVERS>
        <DumpPdb name="dump10" fname="match_rt_under_1.0" tag_time="True"/>
        <DumpPdb name="dump15" fname="match_rt_under_1.5" tag_time="True"/>
        <DumpPdb name="dump20" fname="match_rt_under_2.0" tag_time="True"/>
        <DumpPdb name="dump25" fname="match_rt_under_2.5" tag_time="True"/>
        <DumpPdb name="dump60" fname="match_rt_under_6.0" tag_time="True"/>
        <RemodelMover name="remodel10" fast_disulf="True"
match_rt_limit="1.0" quick_and_dirty="True" bypass_fragments="True"
min_disulfides="1" max_disulfides="1" min_loop="8"/>
        <RemodelMover name="remodel15" fast_disulf="True"
match_rt_limit="1.5" quick_and_dirty="True" bypass_fragments="True"
min_disulfides="1" max_disulfides="1" min_loop="8"/>
        <RemodelMover name="remodel20" fast_disulf="True" match_rt_limit="2"
quick_and_dirty="True" bypass_fragments="True" min_disulfides="1"
max_disulfides="1" min_loop="8"/>
        <RemodelMover name="remodel25" fast_disulf="True"
match_rt_limit="2.5" quick_and_dirty="True" bypass_fragments="True"
min_disulfides="1" max_disulfides="1" min_loop="8"/>
        <RemodelMover name="remodel60" fast_disulf="True" match_rt_limit="1"
quick_and_dirty="True" bypass_fragments="True" min_disulfides="1"
```

```
max_disulfides="1" min_loop="8"/>
        <FastRelax name="relax" task_operations="restrict,cyd_only"
repeats="1"/>
        <PackRotamersMover name="pack" task_operations="cyd_only"/>
        <ParsedProtocol name="build_disulf10">
            <Add mover_name="remodel10"/>
            <Add filter_name="count_cyd"/>
            <Add mover_name="relax"/>
            <Add mover_name="dump10"/>
        </ParsedProtocol>
        <ParsedProtocol name="build_disulf15">
            <Add mover_name="remodel15"/>
            <Add filter_name="count_cyd"/>
            <Add mover_name="relax"/>
            <Add mover_name="dump15"/>
        </ParsedProtocol>
        <ParsedProtocol name="build_disulf20">
            <Add mover_name="remodel20"/>
            <Add filter_name="count_cyd"/>
            <Add mover_name="relax"/>
            <Add mover_name="dump20"/>
        </ParsedProtocol>
        <ParsedProtocol name="build_disulf25">
            <Add mover_name="remodel25"/>
            <Add filter_name="count_cyd"/>
            <Add mover_name="relax"/>
            <Add mover_name="dump25"/>
        </ParsedProtocol>
        <ParsedProtocol name="build_disulf60">
            <Add mover_name="remodel60"/>
            <Add mover_name="relax"/>
            <Add filter_name="count_cyd"/>
            Add filter_name="dslf"/>
            <Add mover_name="dump60"/>
        </ParsedProtocol>
        <LoopOver name="disulfide_loop10" mover_name="build_disulf10"
iterations="10" drift="0"/>
        <LoopOver name="disulfide_loop15" mover_name="build_disulf15"
iterations="10" drift="0"/>
        <LoopOver name="disulfide_loop20" mover_name="build_disulf20"
iterations="10" drift="0"/>
        <LoopOver name="disulfide_loop25" mover_name="build_disulf25"
iterations="10" drift="0"/>
        <LoopOver name="disulfide_loop60" mover_name="build_disulf60"
iterations="25" drift="0"/>
    </MOVERS>
    <PROTOCOLS>
        <Add mover_name="disulfide_loop%%stringency%%"/>
        MAKE SURE TO USE -save_top:1000 at the command line
    </PROTOCOLS>
</dock_design>
```

Protocol S1-4: Representative command and protocol for adding disulfide bonds to a structure. Stringency threshold for rotation-translation matrix is specified at the command line level.

Optimization of Designed Binders Against Human Interleukin-2 Receptor

Primer	**Sequence**
Rbg32.7_SSM_01F	AGGGTCGGCTAGCCATATGNNKGCGCAGGTTATGCTGGA
Rbg32.7_SSM_01R	CATATGGCTAGCCGACCCT
Rbg32.7_SSM_02F	GGTCGGCTAGCCATATGTCCNNKCAGGTTATGCTGGAGGACC
Rbg32.7_SSM_02R	GGACATATGGCTAGCCGACC
Rbg32.7_SSM_03F	GGCTAGCCATATGTCCGCGNNKGTTATGCTGGAGGACCTGG
Rbg32.7_SSM_03R	CGCGGACATATGGCTAGCC
Rbg32.7_SSM_04F	AGCCATATGTCCGCGCAGNNKATGCTGGAGGACCTGGC
Rbg32.7_SSM_04R	CTGCGCGGACATATGGCT
Rbg32.7_SSM_05F	CCATATGTCCGCGCAGGTTNNKCTGGAGGACCTGGCGC
Rbg32.7_SSM_05R	AACCTGCGCGGACATATGG
Rbg32.7_SSM_06F	ATGTCCGCGCAGGTTATGNNKGAGGACCTGGCGCGT
Rbg32.7_SSM_06R	CATAACCTGCGCGGACAT
Rbg32.7_SSM_07F	TCCGCGCAGGTTATGCTGNNKGACCTGGCGCGTGAA
Rbg32.7_SSM_07R	CAGCATAACCTGCGCGGA
Rbg32.7_SSM_08F	CGCGCAGGTTATGCTGGAGNNKCTGGCGCGTGAACTGG
Rbg32.7_SSM_08R	CTCCAGCATAACCTGCGCG
Rbg32.7_SSM_09F	CGCAGGTTATGCTGGAGGACNNKGCGCGTGAACTGGCT
Rbg32.7_SSM_09R	GTCCTCCAGCATAACCTGCG
Rbg32.7_SSM_10F	AGGTTATGCTGGAGGACCTGNNKCGTGAACTGGCTATCGCT
Rbg32.7_SSM_10R	CAGGTCCTCCAGCATAACCT
Rbg32.7_SSM_11F	ATGCTGGAGGACCTGGCGNNKGAACTGGCTATCGCTTGTGT
Rbg32.7_SSM_11R	CGCCAGGTCCTCCAGCAT
Rbg32.7_SSM_12F	CTGGAGGACCTGGCGCGTNNKCTGGCTATCGCTTGTGTTAAAG
Rbg32.7_SSM_12R	ACGCGCCAGGTCCTCCAG
Rbg32.7_SSM_13F	GAGGACCTGGCGCGTGAANNKGCTATCGCTTGTGTTAAAGCT
Rbg32.7_SSM_13R	TTCACGCGCCAGGTCCTC
Rbg32.7_SSM_14F	GACCTGGCGCGTGAACTGNNKATCGCTTGTGTTAAAGCTGATAA
Rbg32.7_SSM_14R	CAGTTCACGCGCCAGGTC
Rbg32.7_SSM_15F	CTGGCGCGTGAACTGGCTNNKGCTTGTGTTAAAGCTGATAAAGAAG
Rbg32.7_SSM_15R	AGCCAGTTCACGCGCCAG
Rbg32.7_SSM_16F	GGCGCGTGAACTGGCTATCNNKTGTGTTAAAGCTGATAAAGAAGGTAA
Rbg32.7_SSM_16R	GATAGCCAGTTCACGCGCC
Rbg32.7_SSM_17F	GCGTGAACTGGCTATCGCTNNKGTTAAAGCTGATAAAGAAGGTAAAGTTG
Rbg32.7_SSM_17R	AGCGATAGCCAGTTCACGC
Rbg32.7_SSM_18F	CGTGAACTGGCTATCGCTTGTNNKAAAGCTGATAAAGAAGGTAAAGTTGA
Rbg32.7_SSM_18R	ACAAGCGATAGCCAGTTCACG
Rbg32.7_SSM_19F	AACTGGCTATCGCTTGTGTTNNKGCTGATAAAGAAGGTAAAGTTGAAGA
Rbg32.7_SSM_19R	AACACAAGCGATAGCCAGTT

Rbg32.7_SSM_20F	AACTGGCTATCGCTTGTGTTAAANNKGATAAAGAAGGTAAAGTTGAAGACGC
Rbg32.7_SSM_20R	TTTAACACAAGCGATAGCCAGTT
Rbg32.7_SSM_21F	GGCTATCGCTTGTGTTAAAGCTNNKAAAGAAGGTAAAGTTGAAGACGC
Rbg32.7_SSM_21R	AGCTTTAACACAAGCGATAGCC
Rbg32.7_SSM_22F	GCTATCGCTTGTGTTAAAGCTGATNNKGAAGGTAAAGTTGAAGACGCTT
Rbg32.7_SSM_22R	ATCAGCTTTAACACAAGCGATAGC
Rbg32.7_SSM_23F	ATCGCTTGTGTTAAAGCTGATAAANNKGGTAAAGTTGAAGACGCTTGC
Rbg32.7_SSM_23R	TTTATCAGCTTTAACACAAGCGAT
Rbg32.7_SSM_24F	TCGCTTGTGTTAAAGCTGATAAAGAANNKAAAGTTGAAGACGCTTGCAC
Rbg32.7_SSM_24R	TTCTTTATCAGCTTTAACACAAGCGA
Rbg32.7_SSM_25F	CTTGTGTTAAAGCTGATAAAGAAGGTNNKGTTGAAGACGCTTGCACC
Rbg32.7_SSM_25R	ACCTTCTTTATCAGCTTTAACACAAG
Rbg32.7_SSM_26F	TTGTGTTAAAGCTGATAAAGAAGGTAAANNKGAAGACGCTTGCACCTATTG
Rbg32.7_SSM_26R	TTTACCTTCTTTATCAGCTTTAACACAA
Rbg32.7_SSM_27F	GTGTTAAAGCTGATAAAGAAGGTAAAGTTNNKGACGCTTGCACCTATTGTG
Rbg32.7_SSM_27R	AACTTTACCTTCTTTATCAGCTTTAACAC
Rbg32.7_SSM_28F	AAAGCTGATAAAGAAGGTAAAGTTGAANNKGCTTGCACCTATTGTGAACAC
Rbg32.7_SSM_28R	TTCAACTTTACCTTCTTTATCAGCTTT
Rbg32.7_SSM_29F	AGCTGATAAAGAAGGTAAAGTTGAAGACNNKTGCACCTATTGTGAACACGC
Rbg32.7_SSM_29R	GTCTTCAACTTTACCTTCTTTATCAGCT
Rbg32.7_SSM_30F	AAGAAGGTAAAGTTGAAGACGCTNNKACCTATTGTGAACACGCGC
Rbg32.7_SSM_30R	AGCGTCTTCAACTTTACCTTCTT
Rbg32.7_SSM_31F	AGGTAAAGTTGAAGACGCTTGCNNKTATTGTGAACACGCGCTGC
Rbg32.7_SSM_31R	GCAAGCGTCTTCAACTTTACCT
Rbg32.7_SSM_32F	AGTTGAAGACGCTTGCACCNNKTGTGAACACGCGCTGC
Rbg32.7_SSM_32R	GGTGCAAGCGTCTTCAACT
Rbg32.7_SSM_33F	GTTGAAGACGCTTGCACCTATNNKGAACACGCGCTGCTCG
Rbg32.7_SSM_33R	ATAGGTGCAAGCGTCTTCAAC
Rbg32.7_SSM_34F	AAGACGCTTGCACCTATTGTNNKCACGCGCTGCTCGA
Rbg32.7_SSM_34R	ACAATAGGTGCAAGCGTCTT
Rbg32.7_SSM_35F	GACGCTTGCACCTATTGTGAANNKGCGCTGCTCGACCTG
Rbg32.7_SSM_35R	TTCACAATAGGTGCAAGCGTC
Rbg32.7_SSM_36F	CGCTTGCACCTATTGTGAACACNNKCTGCTCGACCTGCAGC
Rbg32.7_SSM_36R	GTGTTCACAATAGGTGCAAGCG
Rbg32.7_SSM_37F	GCACCTATTGTGAACACGCGNNKCTCGACCTGCAGCAGATC
Rbg32.7_SSM_37R	CGCGTGTTCACAATAGGTGC
Rbg32.7_SSM_38F	CCTATTGTGAACACGCGCTGNNKGACCTGCAGCAGATCATCG
Rbg32.7_SSM_38R	CAGCGCGTGTTCACAATAGG
Rbg32.7_SSM_39F	TGTGAACACGCGCTGCTCNNKCTGCAGCAGATCATCGTTCT
Rbg32.7_SSM_39R	GAGCAGCGCGTGTTCACA
Rbg32.7_SSM_40F	GAACACGCGCTGCTCGACNNKCAGCAGATCATCGTTCTGTACC

Rbg32.7_SSM_40R	GTCGAGCAGCGCGTGTTC
Rbg32.7_SSM_41F	CACGCGCTGCTCGACCTGNNKCAGATCATCGTTCTGTACCCG
Rbg32.7_SSM_41R	CAGGTCGAGCAGCGCGTG
Rbg32.7_SSM_42F	GCGCTGCTCGACCTGCAGNNKATCATCGTTCTGTACCCGGA
Rbg32.7_SSM_42R	CTGCAGGTCGAGCAGCGC
Rbg32.7_SSM_43F	CTGCTCGACCTGCAGCAGNNKATCGTTCTGTACCCGGAATC
Rbg32.7_SSM_43R	CTGCTGCAGGTCGAGCAG
Rbg32.7_SSM_44F	GCTCGACCTGCAGCAGATCNNKGTTCTGTACCCGGAATCTGT
Rbg32.7_SSM_44R	GATCTGCTGCAGGTCGAGC
Rbg32.7_SSM_45F	TCGACCTGCAGCAGATCATCNNKCTGTACCCGGAATCTGTTGC
Rbg32.7_SSM_45R	GATGATCTGCTGCAGGTCGA
Rbg32.7_SSM_46F	ACCTGCAGCAGATCATCGTTNNKTACCCGGAATCTGTTGCG
Rbg32.7_SSM_46R	AACGATGATCTGCTGCAGGT
Rbg32.7_SSM_47F	CTGCAGCAGATCATCGTTCTGNNKCCGGAATCTGTTGCGCG
Rbg32.7_SSM_47R	CAGAACGATGATCTGCTGCAG
Rbg32.7_SSM_48F	TGCAGCAGATCATCGTTCTGTACNNKGAATCTGTTGCGCGTACC
Rbg32.7_SSM_48R	GTACAGAACGATGATCTGCTGCA
Rbg32.7_SSM_49F	CAGATCATCGTTCTGTACCCGNNKTCTGTTGCGCGTACCG
Rbg32.7_SSM_49R	CGGGTACAGAACGATGATCTG
Rbg32.7_SSM_50F	TCATCGTTCTGTACCCGGAANNKGTTGCGCGTACCGCG
Rbg32.7_SSM_50R	TTCCGGGTACAGAACGATGA
Rbg32.7_SSM_51F	ATCGTTCTGTACCCGGAATCTNNKGCGCGTACCGCGT
Rbg32.7_SSM_51R	AGATTCCGGGTACAGAACGAT
Rbg32.7_SSM_52F	CGTTCTGTACCCGGAATCTGTTNNKCGTACCGCGTACGAACA
Rbg32.7_SSM_52R	AACAGATTCCGGGTACAGAACG
Rbg32.7_SSM_53F	TACCCGGAATCTGTTGCGNNKACCGCGTACGAACAGATG
Rbg32.7_SSM_53R	CGCAACAGATTCCGGGTA
Rbg32.7_SSM_54F	CCGGAATCTGTTGCGCGTNNKGCGTACGAACAGATGATCAC
Rbg32.7_SSM_54R	ACGCGCAACAGATTCCGG
Rbg32.7_SSM_55F	GGAATCTGTTGCGCGTACCNNKTACGAACAGATGATCACCGAA
Rbg32.7_SSM_55R	GGTACGCGCAACAGATTCC
Rbg32.7_SSM_56F	TCTGTTGCGCGTACCGCGNNKGAACAGATGATCACCGAAGCG
Rbg32.7_SSM_56R	CGCGGTACGCGCAACAGA
Rbg32.7_SSM_57F	GTTGCGCGTACCGCGTACNNKCAGATGATCACCGAAGCGC
Rbg32.7_SSM_57R	GTACGCGGTACGCGCAAC
Rbg32.7_SSM_58F	GCGCGTACCGCGTACGAANNKATGATCACCGAAGCGCAG
Rbg32.7_SSM_58R	TTCGTACGCGGTACGCGC
Rbg32.7_SSM_59F	GCGTACCGCGTACGAACAGNNKATCACCGAAGCGCAGC
Rbg32.7_SSM_59R	CTGTTCGTACGCGGTACGC
Rbg32.7_SSM_60F	TACCGCGTACGAACAGATGNNKACCGAAGCGCAGCG
Rbg32.7_SSM_60R	CATCTGTTCGTACGCGGTA

Rbg32.7_SSM_61F	ACCGCGTACGAACAGATGATCNNKGAAGCGCAGCGCCG
Rbg32.7_SSM_61R	GATCATCTGTTCGTACGCGGT
Rbg32.7_SSM_62F	CGCGTACGAACAGATGATCACCNNKGCGCAGCGCCGT
Rbg32.7_SSM_62R	GGTGATCATCTGTTCGTACGCG
Rbg32.7_SSM_63F	TACGAACAGATGATCACCGAANNKCAGCGCCGTATCGCG
Rbg32.7_SSM_63R	TTCGGTGATCATCTGTTCGTA
Rbg32.7_SSM_64F	AACAGATGATCACCGAAGCGNNKCGCCGTATCGCGAACT
Rbg32.7_SSM_64R	CGCTTCGGTGATCATCTGTT
Rbg32.7_SSM_65F	ATGATCACCGAAGCGCAGNNKCGTATCGCGAACTGCGA
Rbg32.7_SSM_65R	CTGCGCTTCGGTGATCAT
Rbg32.7_SSM_66F	ATCACCGAAGCGCAGCGCNNKATCGCGAACTGCGAAAAAG
Rbg32.7_SSM_66R	GCGCTGCGCTTCGGTGAT
Rbg32.7_SSM_67F	ACCGAAGCGCAGCGCCGTNNKGCGAACTGCGAAAAAGTTCT
Rbg32.7_SSM_67R	ACGGCGCTGCGCTTCGGT
Rbg32.7_SSM_68F	GAAGCGCAGCGCCGTATCNNKAACTGCGAAAAAGTTCTGCTC
Rbg32.7_SSM_68R	GATACGGCGCTGCGCTTC
Rbg32.7_SSM_69F	GCGCAGCGCCGTATCGCGNNKTGCGAAAAAGTTCTGCTCG
Rbg32.7_SSM_69R	CGCGATACGGCGCTGCGC
Rbg32.7_SSM_70F	CAGCGCCGTATCGCGAACNNKGAAAAAGTTCTGCTCGAGGG
Rbg32.7_SSM_70R	GTTCGCGATACGGCGCTG
Rbg32.7_SSM_71F	CGCCGTATCGCGAACTGCNNKAAAGTTCTGCTCGAGGGAG
Rbg32.7_SSM_71R	GCAGTTCGCGATACGGCG
Rbg32.7_SSM_72F	CCGTATCGCGAACTGCGAANNKGTTCTGCTCGAGGGAGGC
Rbg32.7_SSM_72R	TTCGCAGTTCGCGATACGG
Rbg32.7_SSM_73F	CGTATCGCGAACTGCGAAAAANNKCTGCTCGAGGGAGGCG
Rbg32.7_SSM_73R	TTTTTCGCAGTTCGCGATACG
Rbg32.7_SSM_74F	TCGCGAACTGCGAAAAAGTTNNKCTCGAGGGAGGCGGAT
Rbg32.7_SSM_74R	AACTTTTTCGCAGTTCGCGA
COF	TGACAACTATATGCGAGCAAATCCCCTCAC
COR	AACTTTTTCGCAGTTCGCGA

Table S2-1: Primers used in the creation of SSM library for Rbg32.7. All sequences are written in 5'→ 3' orientation.

Primer	**Sequence**
Rbg40_01F	AGGGTCGGCTAGCCATATGNNKACCAAAAAACTCCAGCTGCA
Rbg40_01R	CATATGGCTAGCCGACCCT
Rbg40_02F	GGGTCGGCTAGCCATATGTCTNNKAAAAAACTCCAGCTGCAGGC
Rbg40_02R	AGACATATGGCTAGCCGACCC
Rbg40_03F	TCGGCTAGCCATATGTCTACCNNKAAACTCCAGCTGCAGGC
Rbg40_03R	GGTAGACATATGGCTAGCCGA
Rbg40_04F	CGGCTAGCCATATGTCTACCAAANNKCTCCAGCTGCAGGCG
Rbg40_04R	TTTGGTAGACATATGGCTAGCCG

Rbg40_05F	GGCTAGCCATATGTCTACCAAAAAANNKCAGCTGCAGGCGGAAC
Rbg40_05R	TTTTTTGGTAGACATATGGCTAGCC
Rbg40_06F	TAGCCATATGTCTACCAAAAAACTCNNKCTGCAGGCGGAACACG
Rbg40_06R	GAGTTTTTTGGTAGACATATGGCTA
Rbg40_07F	GCCATATGTCTACCAAAAAACTCCAGNNKCAGGCGGAACACGCG
Rbg40_07R	CTGGAGTTTTTTGGTAGACATATGGC
Rbg40_08F	TGTCTACCAAAAAACTCCAGCTGNNKGCGGAACACGCGCT
Rbg40_08R	CAGCTGGAGTTTTTTGGTAGACA
Rbg40_09F	ACCAAAAAACTCCAGCTGCAGNNKGAACACGCGCTCCTGG
Rbg40_09R	CTGCAGCTGGAGTTTTTTGGT
Rbg40_10F	AAACTCCAGCTGCAGGCGNNKCACGCGCTCCTGGAC
Rbg40_10R	CGCCTGCAGCTGGAGTTT
Rbg40_11F	CTCCAGCTGCAGGCGGAANNKGCGCTCCTGGACGC
Rbg40_11R	TTCCGCCTGCAGCTGGAG
Rbg40_12F	CAGCTGCAGGCGGAACACNNKCTCCTGGACGCGCAG
Rbg40_12R	GTGTTCCGCCTGCAGCTG
Rbg40_13F	CTGCAGGCGGAACACGCGNNKCTGGACGCGCAGATGAT
Rbg40_13R	CGCGTGTTCCGCCTGCAG
Rbg40_14F	CAGGCGGAACACGCGCTCNNKGACGCGCAGATGATGCT
Rbg40_14R	GAGCGCGTGTTCCGCCTG
Rbg40_15F	GCGGAACACGCGCTCCTGNNKGCGCAGATGATGCTCAAC
Rbg40_15R	CAGGAGCGCGTGTTCCGC
Rbg40_16F	GAACACGCGCTCCTGGACNNKCAGATGATGCTCAACCGTTC
Rbg40_16R	GTCCAGGAGCGCGTGTTC
Rbg40_17F	CACGCGCTCCTGGACGCGNNKATGATGCTCAACCGTTCTCC
Rbg40_17R	CGCGTCCAGGAGCGCGTG
Rbg40_18F	GCGCTCCTGGACGCGCAGNNKATGCTCAACCGTTCTCCG
Rbg40_18R	CTGCGCGTCCAGGAGCGC
Rbg40_19F	CTCCTGGACGCGCAGATGNNKCTCAACCGTTCTCCGGAAC
Rbg40_19R	CATCTGCGCGTCCAGGAG
Rbg40_20F	CTGGACGCGCAGATGATGNNKAACCGTTCTCCGGAACCG
Rbg40_20R	CATCATCTGCGCGTCCAG
Rbg40_21F	GACGCGCAGATGATGCTCNNKCGTTCTCCGGAACCGAAC
Rbg40_21R	GAGCATCATCTGCGCGTC
Rbg40_22F	CGCGCAGATGATGCTCAACNNKTCTCCGGAACCGAACGA
Rbg40_22R	GTTGAGCATCATCTGCGCG
Rbg40_23F	CGCAGATGATGCTCAACCGTNNKCCGGAACCGAACGAAAAAC
Rbg40_23R	ACGGTTGAGCATCATCTGCG
Rbg40_24F	CAGATGATGCTCAACCGTTCTNNKGAACCGAACGAAAAACTGAACC
Rbg40_24R	AGAACGGTTGAGCATCATCTG
Rbg40_25F	ATGCTCAACCGTTCTCCGNNKCCGAACGAAAAACTGAACCG

Rbg40_25R	CGGAGAACGGTTGAGCAT
Rbg40_26F	TGCTCAACCGTTCTCCGGAANNKAACGAAAAACTGAACCGTATCAT
Rbg40_26R	TTCCGGAGAACGGTTGAGCA
Rbg40_27F	AACCGTTCTCCGGAACCGNNKGAAAAACTGAACCGTATCATCACC
Rbg40_27R	CGGTTCCGGAGAACGGTT
Rbg40_28F	CCGTTCTCCGGAACCGAACNNKAAACTGAACCGTATCATCACCA
Rbg40_28R	GTTCGGTTCCGGAGAACGG
Rbg40_29F	TTCTCCGGAACCGAACGAANNKCTGAACCGTATCATCACCACC
Rbg40_29R	TTCGTTCGGTTCCGGAGAA
Rbg40_30F	CTCCGGAACCGAACGAAAAANNKAACCGTATCATCACCACCATG
Rbg40_30R	TTTTTCGTTCGGTTCCGGAG
Rbg40_31F	CCGGAACCGAACGAAAAACTGNNKCGTATCATCACCACCATGCA
Rbg40_31R	CAGTTTTTCGTTCGGTTCCGG
Rbg40_32F	CGGAACCGAACGAAAAACTGAACNNKATCATCACCACCATGCAGTC
Rbg40_32R	GTTCAGTTTTTCGTTCGGTTCCG
Rbg40_33F	CCGAACGAAAAACTGAACCGTNNKATCACCACCATGCAGTCTTG
Rbg40_33R	ACGGTTCAGTTTTTCGTTCGG
Rbg40_34F	CCGAACGAAAAACTGAACCGTATCNNKACCACCATGCAGTCTTGG
Rbg40_34R	GATACGGTTCAGTTTTTCGTTCGG
Rbg40_35F	AACGAAAAACTGAACCGTATCATCNNKACCATGCAGTCTTGGATCTC
Rbg40_35R	GATGATACGGTTCAGTTTTTCGTT
Rbg40_36F	GAAAAACTGAACCGTATCATCACCNNKATGCAGTCTTGGATCTCTACTG
Rbg40_36R	GGTGATGATACGGTTCAGTTTTTC
Rbg40_37F	CTGAACCGTATCATCACCACCNNKCAGTCTTGGATCTCTACTGGTAAG
Rbg40_37R	GGTGGTGATGATACGGTTCAG
Rbg40_38F	AACCGTATCATCACCACCATGNNKTCTTGGATCTCTACTGGTAAGATCG
Rbg40_38R	CATGGTGGTGATGATACGGTT
Rbg40_39F	CCGTATCATCACCACCATGCAGNNKTGGATCTCTACTGGTAAGATCGA
Rbg40_39R	CTGCATGGTGGTGATGATACGG
Rbg40_40F	TCATCACCACCATGCAGTCTNNKATCTCTACTGGTAAGATCGACCT
Rbg40_40R	AGACTGCATGGTGGTGATGA
Rbg40_41F	CACCACCATGCAGTCTTGGNNKTCTACTGGTAAGATCGACCTGG
Rbg40_41R	CCAAGACTGCATGGTGGTG
Rbg40_42F	ACCACCATGCAGTCTTGGATCNNKACTGGTAAGATCGACCTGGA
Rbg40_42R	GATCCAAGACTGCATGGTGGT
Rbg40_43F	ACCATGCAGTCTTGGATCTCTNNKGGTAAGATCGACCTGGATGG
Rbg40_43R	AGAGATCCAAGACTGCATGGT
Rbg40_44F	CCATGCAGTCTTGGATCTCTACTNNKAAGATCGACCTGGATGGTGC
Rbg40_44R	AGTAGAGATCCAAGACTGCATGG
Rbg40_45F	TGCAGTCTTGGATCTCTACTGGTNNKATCGACCTGGATGGTGCG
Rbg40_45R	ACCAGTAGAGATCCAAGACTGCA

Rbg40_46F	GCAGTCTTGGATCTCTACTGGTAAGNNKGACCTGGATGGTGCGAAG
Rbg40_46R	CTTACCAGTAGAGATCCAAGACTGC
Rbg40_47F	AGTCTTGGATCTCTACTGGTAAGATCNNKCTGGATGGTGCGAAGGAA
Rbg40_47R	GATCTTACCAGTAGAGATCCAAGACT
Rbg40_48F	TGGATCTCTACTGGTAAGATCGACNNKGATGGTGCGAAGGAACTCG
Rbg40_48R	GTCGATCTTACCAGTAGAGATCCA
Rbg40_49F	TCTCTACTGGTAAGATCGACCTGNNKGGTGCGAAGGAACTCGC
Rbg40_49R	CAGGTCGATCTTACCAGTAGAGA
Rbg40_50F	TACTGGTAAGATCGACCTGGATNNKGCGAAGGAACTCGCGA
Rbg40_50R	ATCCAGGTCGATCTTACCAGTA
Rbg40_51F	TGGTAAGATCGACCTGGATGGTNNKAAGGAACTCGCGAAAGAAGT
Rbg40_51R	ACCATCCAGGTCGATCTTACCA
Rbg40_52F	ATCGACCTGGATGGTGCGNNKGAACTCGCGAAAGAAGTTGAA
Rbg40_52R	CGCACCATCCAGGTCGAT
Rbg40_53F	CGACCTGGATGGTGCGAAGNNKCTCGCGAAAGAAGTTGAAGAA
Rbg40_53R	CTTCGCACCATCCAGGTCG
Rbg40_54F	CCTGGATGGTGCGAAGGAANNKGCGAAAGAAGTTGAAGAACTGC
Rbg40_54R	TTCCTTCGCACCATCCAGG
Rbg40_55F	TGGATGGTGCGAAGGAACTCNNKAAAGAAGTTGAAGAACTGCGTC
Rbg40_55R	GAGTTCCTTCGCACCATCCA
Rbg40_56F	GGTGCGAAGGAACTCGCGNNKGAAGTTGAAGAACTGCGTCAG
Rbg40_56R	CGCGAGTTCCTTCGCACC
Rbg40_57F	TGCGAAGGAACTCGCGAAANNKGTTGAAGAACTGCGTCAGGA
Rbg40_57R	TTTCGCGAGTTCCTTCGCA
Rbg40_58F	GCGAAGGAACTCGCGAAAGAANNKGAAGAACTGCGTCAGGAAGC
Rbg40_58R	TTCTTTCGCGAGTTCCTTCGC
Rbg40_59F	AAGGAACTCGCGAAAGAAGTTNNKGAACTGCGTCAGGAAGCG
Rbg40_59R	AACTTCTTTCGCGAGTTCCTT
Rbg40_60F	GGAACTCGCGAAAGAAGTTGAANNKCTGCGTCAGGAAGCGG
Rbg40_60R	TTCAACTTCTTTCGCGAGTTCC
Rbg40_61F	ACTCGCGAAAGAAGTTGAAGAANNKCGTCAGGAAGCGGAAAAAC
Rbg40_61R	TTCTTCAACTTCTTTCGCGAGT
Rbg40_62F	TCGCGAAAGAAGTTGAAGAACTGNNKCAGGAAGCGGAAAAACGTG
Rbg40_62R	CAGTTCTTCAACTTCTTTCGCGA
Rbg40_63F	CGAAAGAAGTTGAAGAACTGCGTNNKGAAGCGGAAAAACGTGGTAT
Rbg40_63R	ACGCAGTTCTTCAACTTCTTTCG
Rbg40_64F	AGAAGTTGAAGAACTGCGTCAGNNKGCGGAAAAACGTGGTATCG
Rbg40_64R	CTGACGCAGTTCTTCAACTTCT
Rbg40_65F	GTTGAAGAACTGCGTCAGGAANNKGAAAAACGTGGTATCGACGTT
Rbg40_65R	TTCCTGACGCAGTTCTTCAAC
Rbg40_66F	GAACTGCGTCAGGAAGCGNNKAAACGTGGTATCGACGTTCG

Rbg40_66R	CGCTTCCTGACGCAGTTC
Rbg40_67F	CTGCGTCAGGAAGCGGAANNKCGTGGTATCGACGTTCGT
Rbg40_67R	TTCCGCTTCCTGACGCAG
Rbg40_68F	TGCGTCAGGAAGCGGAAAAANNKGGTATCGACGTTCGTGACC
Rbg40_68R	TTTTTCCGCTTCCTGACGCA
Rbg40_69F	GTCAGGAAGCGGAAAAACGTNNKATCGACGTTCGTGACCTG
Rbg40_69R	ACGTTTTTCCGCTTCCTGAC
Rbg40_70F	AGGAAGCGGAAAAACGTGGTNNKGACGTTCGTGACCTGGC
Rbg40_70R	ACCACGTTTTTCCGCTTCCT
Rbg40_71F	GAAGCGGAAAAACGTGGTATCNNKGTTCGTGACCTGGCGT
Rbg40_71R	GATACCACGTTTTTCCGCTTC
Rbg40_72F	GCGGAAAAACGTGGTATCGACNNKCGTGACCTGGCGTCTAAC
Rbg40_72R	GTCGATACCACGTTTTTCCGC
Rbg40_73F	GGAAAAACGTGGTATCGACGTTNNKGACCTGGCGTCTAACCTG
Rbg40_73R	AACGTCGATACCACGTTTTTCC
Rbg40_74F	ACGTGGTATCGACGTTCGTNNKCTGGCGTCTAACCTGAAAGT
Rbg40_74R	ACGAACGTCGATACCACGT
Rbg40_75F	GTGGTATCGACGTTCGTGACNNKGCGTCTAACCTGAAAGTTATCC
Rbg40_75R	GTCACGAACGTCGATACCAC
Rbg40_76F	ATCGACGTTCGTGACCTGNNKTCTAACCTGAAAGTTATCCTGCTG
Rbg40_76R	CAGGTCACGAACGTCGAT
Rbg40_77F	GACGTTCGTGACCTGGCGNNKAACCTGAAAGTTATCCTGCTGG
Rbg40_77R	CGCCAGGTCACGAACGTC
Rbg40_78F	GTTCGTGACCTGGCGTCTNNKCTGAAAGTTATCCTGCTGGAGC
Rbg40_78R	AGACGCCAGGTCACGAAC
Rbg40_79F	TCGTGACCTGGCGTCTAACNNKAAAGTTATCCTGCTGGAGCTG
Rbg40_79R	GTTAGACGCCAGGTCACGA
Rbg40_80F	TGACCTGGCGTCTAACCTGNNKGTTATCCTGCTGGAGCTGG
Rbg40_80R	CAGGTTAGACGCCAGGTCA
Rbg40_81F	ACCTGGCGTCTAACCTGAAANNKATCCTGCTGGAGCTGGC
Rbg40_81R	TTTCAGGTTAGACGCCAGGT
Rbg40_82F	CTGGCGTCTAACCTGAAAGTTNNKCTGCTGGAGCTGGCTC
Rbg40_82R	AACTTTCAGGTTAGACGCCAG
Rbg40_83F	TGGCGTCTAACCTGAAAGTTATCNNKCTGGAGCTGGCTCTCGA
Rbg40_83R	GATAACTTTCAGGTTAGACGCCA
Rbg40_84F	GGCGTCTAACCTGAAAGTTATCCTGNNKGAGCTGGCTCTCGAGGG
Rbg40_84R	CAGGATAACTTTCAGGTTAGACGCC
Rbg40_85F	TCTAACCTGAAAGTTATCCTGCTGNNKCTGGCTCTCGAGGGAGG
Rbg40_85R	CAGCAGGATAACTTTCAGGTTAGA
Rbg40_86F	ACCTGAAAGTTATCCTGCTGGAGNNKGCTCTCGAGGGAGGCG
Rbg40_86R	CTCCAGCAGGATAACTTTCAGGT

Rbg40_87F	AAAGTTATCCTGCTGGAGCTGNNKCTCGAGGGAGGCGGAT
Rbg40_87R	CAGCTCCAGCAGGATAACTTT
COF	TGACAACTATATGCGAGCAAATCCCCTCAC
COR	AACTTTTTCGCAGTTCGCGA

Table S2-2: Primers used in the creation of SSM library for Rbg40. All sequences are written in 5'→ 3' orientation.

Oligonucleotide	**Sequence**
Rbg32.7_combinatorial_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCCGCGCAGGTT
Rbg32.7_combinatorial_r2	CAGTTCACGCGCATTGTCCTCCAGCWYAACCTGCGCGGACATATGGC
Rbg32.7_combinatorial_f3	AGGACAATGCGCGTGAACTGYWTATCGCTTGTKKGAAAGCTSWTAAAGWAGGTAAAGTTGAAGACGCTT
Rbg32.7_combinatorial_r4	CTGCTGCAGGTCGAGCAGCSMGTGTTCACAAWWGGTGCAAGCGTCTTCAACTTTACCT
Rbg32.7_combinatorial_f5	GCTGCTCGACCTGCAGCAGMTCMWAGTTCTGTACCCGGAATCTG
Rbg32.7_combinatorial_r6	CGGTGATCATCTGTTCGTAGKCGGTACGCGCAACAGATTCCGGGTACAGAACT
Rbg32.7_combinatorial_f7	CTACGAACAGATGATCACCGAADYGCAGCGCCGTATCKYCAACTGCGAAAAAGTTCTGC
Rbg32.7_combinatorial_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGCAGAACTTTTTCGCAGTTG
Rbg40_combinatorial_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCTACCAAAAAAACCCAGCTG
Rbg40_combinatorial_r2	GCGTCCAGGAGCGCGTGTTCCGCSABCAGCTGGGTTTTTTTGGTAG
Rbg40_combinatorial_f3	GAACACGCGCTCCTGGACGCGYWCATGATGCTCAACDTKDTGCCGGAACCGAACGAAAAA
Rbg40_combinatorial_r4	GATCCAAGACTGCATGGTGGTGATGATACGGTTCAGTTTTTCGTTCGGTTCCGGCA
Rbg40_combinatorial_f5	CCACCATGCAGTCTTGGATCYWTACTGGTAAGATCGACSSGGATGGTGCGMAGGAACTCGCGAAAGAAGTTG
Rbg40_combinatorial_r6	ACGTCGATACCACGTTTTTCGWRTTCCTGWTCCAGTTCTTCAACTTCTTTCGCGAGTTCC
Rbg40_combinatorial_f7	GAAAAACGTGGTATCGACGTTSAKGACDASGCGTCTAACCTGAAAGTTATCCTGCT
Rbg40_combinatorial_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGAGCCAGCTCCAGCAGGATAACTTTCAGGTT

Table S2-3: Oligonucleotides used in the assembly of combinatorial libraries based on Rbg32.7 and Rbg40. All sequences are written in 5'→ 3' orientation.

Primer	**Templates**	**Sequence**
5'_Rbg32.8	Rbg32.8A to Rbg32.8L	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCCGCGCAGGTTG
3'_Rbg32.8	Rbg32.8A to Rbg32.8H, Rbg32.8J to Rbg32.8L	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCAGAACTTTTTCGC
3'_Rbg32.8I	Rbg32.8	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCAGAAATTTTTCGC
5'_Rbg40.1	Rbg40.1A to Rbg40.1G, Rbg40.1I, Rbg40.1J	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAAAAACCCAG
5'_Rbg40.1H	Rbg40.1H	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAAAAACCCAC

3'_Rbg40.1	Rbg40.1A to Rbg40.1J	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAG CCAGCTCCAG

Table S2-4: Primers used to subclone genes encoding second- and third-generation binders against hIL2β/γ_c into pET29b. All sequences are written in 5'→ 3' orientation.

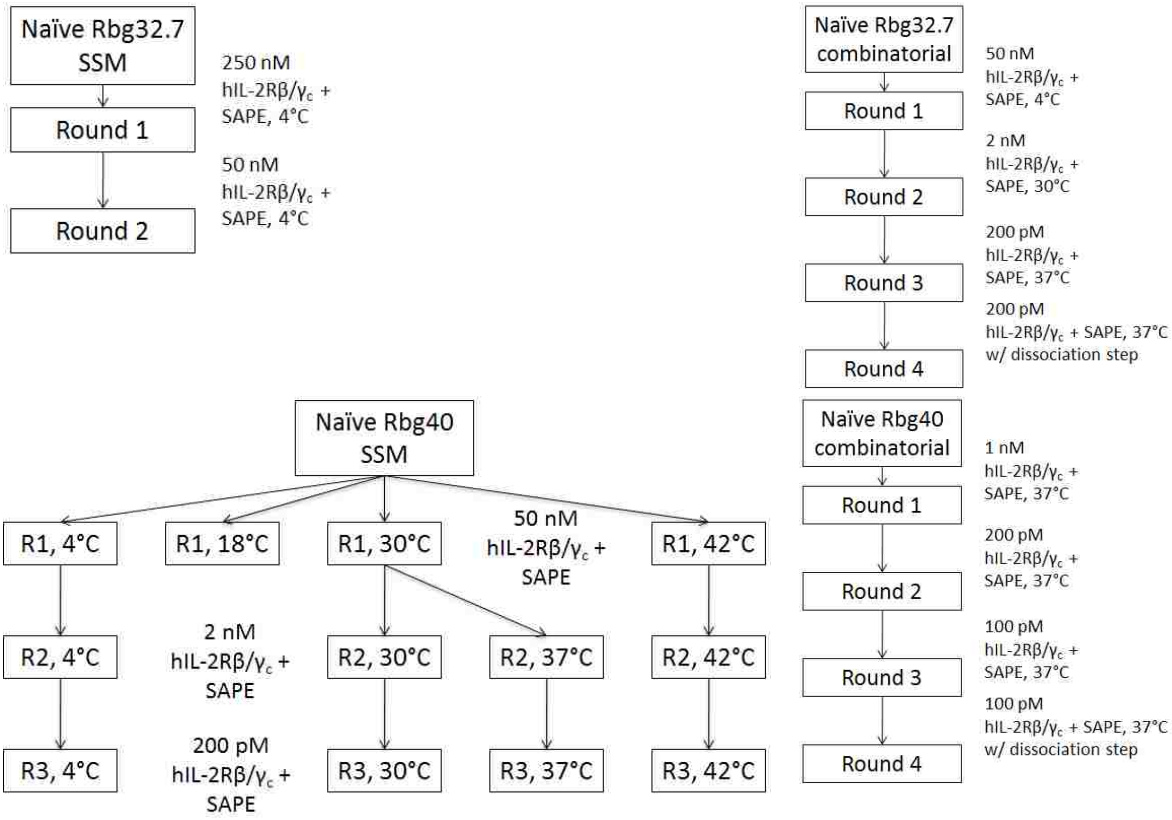


Figure S2-1: Selection schemes for Rbg32.7 and Rbg40 SSM and combinatorial libraries.

Optimization of Designed Binders Against Mouse Interleukin-2 Receptor

Primer	**Sequence**
Rbg40.1F_01F	AGGGTCGGCTAGCCATATGNNKACCAAAAAAACCCAGCTGC
Rbg40.1F_01R	CATATGGCTAGCCGACCCT
Rbg40.1F_02F	GGGTCGGCTAGCCATATGTCTNNKAAAAAAACCCAGCTGCTGG
Rbg40.1F_02R	AGACATATGGCTAGCCGACCC
Rbg40.1F_03F	TCGGCTAGCCATATGTCTACCNNKAAAACCCAGCTGCTGGC
Rbg40.1F_03R	GGTAGACATATGGCTAGCCGA
Rbg40.1F_04F	CGGCTAGCCATATGTCTACCAAANNKACCCAGCTGCTGGCG
Rbg40.1F_04R	TTTGGTAGACATATGGCTAGCCG
Rbg40.1F_05F	GGCTAGCCATATGTCTACCAAAAAANNKCAGCTGCTGGCGGAAC
Rbg40.1F_05R	TTTTTTGGTAGACATATGGCTAGCC
Rbg40.1F_06F	TAGCCATATGTCTACCAAAAAAACCNNKCTGCTGGCGGAACACG
Rbg40.1F_06R	GGTTTTTTTGGTAGACATATGGCTA
Rbg40.1F_07F	GCCATATGTCTACCAAAAAAACCCAGNNKCTGGCGGAACACGCG

Rbg40.1F_07R	CTGGGTTTTTTTGGTAGACATATGGC
Rbg40.1F_08F	TGTCTACCAAAAAAACCCAGCTGNNKGCGGAACACGCGCT
Rbg40.1F_08R	CAGCTGGGTTTTTTTGGTAGACA
Rbg40.1F_09F	ACCAAAAAAACCCAGCTGCTGNNKGAACACGCGCTCCTGG
Rbg40.1F_09R	CAGCAGCTGGGTTTTTTTGGT
Rbg40.1F_10F	AAAACCCAGCTGCTGGCGNNKCACGCGCTCCTGGAC
Rbg40.1F_10R	CGCCAGCAGCTGGGTTTT
Rbg40.1F_11F	ACCCAGCTGCTGGCGGAANNKGCGCTCCTGGACGC
Rbg40.1F_11R	TTCCGCCAGCAGCTGGGT
Rbg40.1F_12F	CAGCTGCTGGCGGAACACNNKCTCCTGGACGCGCTCA
Rbg40.1F_12R	GTGTTCCGCCAGCAGCTG
Rbg40.1F_13F	CTGCTGGCGGAACACGCGNNKCTGGACGCGCTCATGAT
Rbg40.1F_13R	CGCGTGTTCCGCCAGCAG
Rbg40.1F_14F	CTGGCGGAACACGCGCTCNNKGACGCGCTCATGATGCT
Rbg40.1F_14R	GAGCGCGTGTTCCGCCAG
Rbg40.1F_15F	GCGGAACACGCGCTCCTGNNKGCGCTCATGATGCTCAAC
Rbg40.1F_15R	CAGGAGCGCGTGTTCCGC
Rbg40.1F_16F	GAACACGCGCTCCTGGACNNKCTCATGATGCTCAACTTGTTGC
Rbg40.1F_16R	GTCCAGGAGCGCGTGTTC
Rbg40.1F_17F	CACGCGCTCCTGGACGCGNNKATGATGCTCAACTTGTTGCC
Rbg40.1F_17R	CGCGTCCAGGAGCGCGTG
Rbg40.1F_18F	GCGCTCCTGGACGCGCTCNNKATGCTCAACTTGTTGCCGG
Rbg40.1F_18R	GAGCGCGTCCAGGAGCGC
Rbg40.1F_19F	CTCCTGGACGCGCTCATGNNKCTCAACTTGTTGCCGGAAC
Rbg40.1F_19R	CATGAGCGCGTCCAGGAG
Rbg40.1F_20F	CTGGACGCGCTCATGATGNNKAACTTGTTGCCGGAACCG
Rbg40.1F_20R	CATCATGAGCGCGTCCAG
Rbg40.1F_21F	GACGCGCTCATGATGCTCNNKTTGTTGCCGGAACCGAA
Rbg40.1F_21R	GAGCATCATGAGCGCGTC
Rbg40.1F_22F	CGCGCTCATGATGCTCAACNNKTTGCCGGAACCGAACG
Rbg40.1F_22R	GTTGAGCATCATGAGCGCG
Rbg40.1F_23F	GCGCTCATGATGCTCAACTTGNNKCCGGAACCGAACGAAAAAC
Rbg40.1F_23R	CAAGTTGAGCATCATGAGCGC
Rbg40.1F_24F	CGCTCATGATGCTCAACTTGTTGNNKGAACCGAACGAAAAACTGAACC
Rbg40.1F_24R	CAACAAGTTGAGCATCATGAGCG
Rbg40.1F_25F	TGATGCTCAACTTGTTGCCGNNKCCGAACGAAAAACTGAACCG
Rbg40.1F_25R	CGGCAACAAGTTGAGCATCA
Rbg40.1F_26F	TGCTCAACTTGTTGCCGGAANNKAACGAAAAACTGAACCGTATCAT
Rbg40.1F_26R	TTCCGGCAACAAGTTGAGCA
Rbg40.1F_27F	AACTTGTTGCCGGAACCGNNKGAAAAACTGAACCGTATCATCACC
Rbg40.1F_27R	CGGTTCCGGCAACAAGTT

Rbg40.1F_28F	TTGTTGCCGGAACCGAACNNKAAACTGAACCGTATCATCACCA
Rbg40.1F_28R	GTTCGGTTCCGGCAACAA
Rbg40.1F_29F	TTGCCGGAACCGAACGAANNKCTGAACCGTATCATCACCACC
Rbg40.1F_29R	TTCGTTCGGTTCCGGCAA
Rbg40.1F_30F	TGCCGGAACCGAACGAAAAANNKAACCGTATCATCACCACCATG
Rbg40.1F_30R	TTTTTCGTTCGGTTCCGGCA
Rbg40.1F_31F	CCGGAACCGAACGAAAAACTGNNKCGTATCATCACCACCATGCA
Rbg40.1F_31R	CAGTTTTTCGTTCGGTTCCGG
Rbg40.1F_32F	CGGAACCGAACGAAAAACTGAACNNKATCATCACCACCATGCAGTC
Rbg40.1F_32R	GTTCAGTTTTTCGTTCGGTTCCG
Rbg40.1F_33F	CCGAACGAAAAACTGAACCGTNNKATCACCACCATGCAGTCTTG
Rbg40.1F_33R	ACGGTTCAGTTTTTCGTTCGG
Rbg40.1F_34F	CCGAACGAAAAACTGAACCGTATCNNKACCACCATGCAGTCTTGG
Rbg40.1F_34R	GATACGGTTCAGTTTTTCGTTCGG
Rbg40.1F_35F	AACGAAAAACTGAACCGTATCATCNNKACCATGCAGTCTTGGATCTTT
Rbg40.1F_35R	GATGATACGGTTCAGTTTTTCGTT
Rbg40.1F_36F	GAAAAACTGAACCGTATCATCACCNNKATGCAGTCTTGGATCTTTACTGG
Rbg40.1F_36R	GGTGATGATACGGTTCAGTTTTTC
Rbg40.1F_37F	CTGAACCGTATCATCACCACCNNKCAGTCTTGGATCTTTACTGGTAAGA
Rbg40.1F_37R	GGTGGTGATGATACGGTTCAG
Rbg40.1F_38F	AACCGTATCATCACCACCATGNNKTCTTGGATCTTTACTGGTAAGATCG
Rbg40.1F_38R	CATGGTGGTGATGATACGGTT
Rbg40.1F_39F	CCGTATCATCACCACCATGCAGNNKTGGATCTTTACTGGTAAGATCGAC
Rbg40.1F_39R	CTGCATGGTGGTGATGATACGG
Rbg40.1F_40F	TCATCACCACCATGCAGTCTNNKATCTTTACTGGTAAGATCGACGG
Rbg40.1F_40R	AGACTGCATGGTGGTGATGA
Rbg40.1F_41F	CACCACCATGCAGTCTTGGNNKTTTACTGGTAAGATCGACGGG
Rbg40.1F_41R	CCAAGACTGCATGGTGGTG
Rbg40.1F_42F	ACCACCATGCAGTCTTGGATCNNKACTGGTAAGATCGACGGGG
Rbg40.1F_42R	GATCCAAGACTGCATGGTGGT
Rbg40.1F_43F	CACCATGCAGTCTTGGATCTTTNNKGGTAAGATCGACGGGGATG
Rbg40.1F_43R	AAAGATCCAAGACTGCATGGTG
Rbg40.1F_44F	ACCATGCAGTCTTGGATCTTTACTNNKAAGATCGACGGGGATGGT
Rbg40.1F_44R	AGTAAAGATCCAAGACTGCATGGT
Rbg40.1F_45F	TGCAGTCTTGGATCTTTACTGGTNNKATCGACGGGGATGGTGC
Rbg40.1F_45R	ACCAGTAAAGATCCAAGACTGCA
Rbg40.1F_46F	TGCAGTCTTGGATCTTTACTGGTAAGNNKGACGGGGATGGTGCG
Rbg40.1F_46R	CTTACCAGTAAAGATCCAAGACTGCA
Rbg40.1F_47F	AGTCTTGGATCTTTACTGGTAAGATCNNKGGGGATGGTGCGCAG
Rbg40.1F_47R	GATCTTACCAGTAAAGATCCAAGACT
Rbg40.1F_48F	TTGGATCTTTACTGGTAAGATCGACNNKGATGGTGCGCAGGAACT

Rbg40.1F_48R	GTCGATCTTACCAGTAAAGATCCAA
Rbg40.1F_49F	CTTTACTGGTAAGATCGACGGGNNKGGTGCGCAGGAACTCG
Rbg40.1F_49R	CCCGTCGATCTTACCAGTAAAG
Rbg40.1F_50F	CTGGTAAGATCGACGGGGATNNKGCGCAGGAACTCGCG
Rbg40.1F_50R	ATCCCCGTCGATCTTACCAG
Rbg40.1F_51F	TAAGATCGACGGGGATGGTNNKCAGGAACTCGCGAAAGAAGT
Rbg40.1F_51R	ACCATCCCCGTCGATCTTA
Rbg40.1F_52F	ATCGACGGGGATGGTGCGNNKGAACTCGCGAAAGAAGTTGAA
Rbg40.1F_52R	CGCACCATCCCCGTCGAT
Rbg40.1F_53F	GACGGGGATGGTGCGCAGNNKCTCGCGAAAGAAGTTGAAGAA
Rbg40.1F_53R	CTGCGCACCATCCCCGTC
Rbg40.1F_54F	GGGGATGGTGCGCAGGAANNKGCGAAAGAAGTTGAAGAACTGG
Rbg40.1F_54R	TTCCTGCGCACCATCCCC
Rbg40.1F_55F	GATGGTGCGCAGGAACTCNNKAAAGAAGTTGAAGAACTGGAACAG
Rbg40.1F_55R	GAGTTCCTGCGCACCATC
Rbg40.1F_56F	GGTGCGCAGGAACTCGCGNNKGAAGTTGAAGAACTGGAACAGG
Rbg40.1F_56R	CGCGAGTTCCTGCGCACC
Rbg40.1F_57F	GCGCAGGAACTCGCGAAANNKGTTGAAGAACTGGAACAGGAAC
Rbg40.1F_57R	TTTCGCGAGTTCCTGCGC
Rbg40.1F_58F	CGCAGGAACTCGCGAAAGAANNKGAAGAACTGGAACAGGAACAC
Rbg40.1F_58R	TTCTTTCGCGAGTTCCTGCG
Rbg40.1F_59F	CAGGAACTCGCGAAAGAAGTTNNKGAACTGGAACAGGAACACGA
Rbg40.1F_59R	AACTTCTTTCGCGAGTTCCTG
Rbg40.1F_60F	GGAACTCGCGAAAGAAGTTGAANNKCTGGAACAGGAACACGAAAAA
Rbg40.1F_60R	TTCAACTTCTTTCGCGAGTTCC
Rbg40.1F_61F	ACTCGCGAAAGAAGTTGAAGAANNKGAACAGGAACACGAAAAACGT
Rbg40.1F_61R	TTCTTCAACTTCTTTCGCGAGT
Rbg40.1F_62F	TCGCGAAAGAAGTTGAAGAACTGNNKCAGGAACACGAAAAACGTGG
Rbg40.1F_62R	CAGTTCTTCAACTTCTTTCGCGA
Rbg40.1F_63F	GCGAAAGAAGTTGAAGAACTGGAANNKGAACACGAAAAACGTGGTATCG
Rbg40.1F_63R	TTCCAGTTCTTCAACTTCTTTCGC
Rbg40.1F_64F	AAAGAAGTTGAAGAACTGGAACAGNNKCACGAAAAACGTGGTATCGAC
Rbg40.1F_64R	CTGTTCCAGTTCTTCAACTTCTTT
Rbg40.1F_65F	AAGTTGAAGAACTGGAACAGGAANNKGAAAAACGTGGTATCGACGTT
Rbg40.1F_65R	TTCCTGTTCCAGTTCTTCAACTT
Rbg40.1F_66F	TGAAGAACTGGAACAGGAACACNNKAAACGTGGTATCGACGTTGA
Rbg40.1F_66R	GTGTTCCTGTTCCAGTTCTTCA
Rbg40.1F_67F	GAACTGGAACAGGAACACGAANNKCGTGGTATCGACGTTGAGG
Rbg40.1F_67R	TTCGTGTTCCTGTTCCAGTTC
Rbg40.1F_68F	ACTGGAACAGGAACACGAAAAANNKGGTATCGACGTTGAGGACTAC
Rbg40.1F_68R	TTTTTCGTGTTCCTGTTCCAGT

Rbg40.1F_69F	GGAACAGGAACACGAAAAACGTNNKATCGACGTTGAGGACTACG
Rbg40.1F_69R	ACGTTTTTCGTGTTCCTGTTCC
Rbg40.1F_70F	AGGAACACGAAAAACGTGGTNNKGACGTTGAGGACTACGCG
Rbg40.1F_70R	ACCACGTTTTTCGTGTTCCT
Rbg40.1F_71F	AGGAACACGAAAAACGTGGTATCNNKGTTGAGGACTACGCGTCTAA
Rbg40.1F_71R	GATACCACGTTTTTCGTGTTCCT
Rbg40.1F_72F	ACACGAAAAACGTGGTATCGACNNKGAGGACTACGCGTCTAACC
Rbg40.1F_72R	GTCGATACCACGTTTTTCGTGT
Rbg40.1F_73F	CGAAAAACGTGGTATCGACGTTNNKGACTACGCGTCTAACCTGAA
Rbg40.1F_73R	AACGTCGATACCACGTTTTTCG
Rbg40.1F_74F	AACGTGGTATCGACGTTGAGNNKTACGCGTCTAACCTGAAAGTT
Rbg40.1F_74R	CTCAACGTCGATACCACGTT
Rbg40.1F_75F	CGTGGTATCGACGTTGAGGACNNKGCGTCTAACCTGAAAGTTATCC
Rbg40.1F_75R	GTCCTCAACGTCGATACCACG
Rbg40.1F_76F	TGGTATCGACGTTGAGGACTACNNKTCTAACCTGAAAGTTATCCTGCTG
Rbg40.1F_76R	GTAGTCCTCAACGTCGATACCA
Rbg40.1F_77F	GACGTTGAGGACTACGCGNNKAACCTGAAAGTTATCCTGCTGG
Rbg40.1F_77R	CGCGTAGTCCTCAACGTC
Rbg40.1F_78F	ACGTTGAGGACTACGCGTCTNNKCTGAAAGTTATCCTGCTGGAGC
Rbg40.1F_78R	AGACGCGTAGTCCTCAACGT
Rbg40.1F_79F	GTTGAGGACTACGCGTCTAACNNKAAAGTTATCCTGCTGGAGCTG
Rbg40.1F_79R	GTTAGACGCGTAGTCCTCAAC
Rbg40.1F_80F	AGGACTACGCGTCTAACCTGNNKGTTATCCTGCTGGAGCTGG
Rbg40.1F_80R	CAGGTTAGACGCGTAGTCCT
Rbg40.1F_81F	GGACTACGCGTCTAACCTGAAANNKATCCTGCTGGAGCTGGC
Rbg40.1F_81R	TTTCAGGTTAGACGCGTAGTCC
Rbg40.1F_82F	TACGCGTCTAACCTGAAAGTTNNKCTGCTGGAGCTGGCTC
Rbg40.1F_82R	AACTTTCAGGTTAGACGCGTA
Rbg40.1F_83F	ACGCGTCTAACCTGAAAGTTATCNNKCTGGAGCTGGCTCTCGA
Rbg40.1F_83R	GATAACTTTCAGGTTAGACGCGT
Rbg40.1F_84F	CGCGTCTAACCTGAAAGTTATCCTGNNKGAGCTGGCTCTCGAGGG
Rbg40.1F_84R	CAGGATAACTTTCAGGTTAGACGCG
Rbg40.1F_85F	TCTAACCTGAAAGTTATCCTGCTGNNKCTGGCTCTCGAGGGAGG
Rbg40.1F_85R	CAGCAGGATAACTTTCAGGTTAGA
Rbg40.1F_86F	ACCTGAAAGTTATCCTGCTGGAGNNKGCTCTCGAGGGAGGCG
Rbg40.1F_86R	CTCCAGCAGGATAACTTTCAGGT
Rbg40.1F_87F	AAAGTTATCCTGCTGGAGCTGNNKCTCGAGGGAGGCGGAT
Rbg40.1F_87R	CAGCTCCAGCAGGATAACTTT
COF	TGACAACTATATGCGAGCAAATCCCCTCAC
COR	AACTTTTTCGCAGTTCGCGA

Table S3-1: Primers used in the creation of SSM library for Rbg40.1F. All sequences are written in 5'→ 3' orientation.

Oligonucleotide	Sequence
Rbg40.1F_mouse_combo_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCTACC
Rbg40.1F_mouse_combo_r2	GGAGCGCGTSTTCCGCCGCGWGCTGGGTWTYTKYGGTAGACATATGGCTAGCCG
Rbg40.1F_mouse_combo_f3	CGCGGCGGAASACGCGCTCCKGGACGCGCTCATGMTGARGMACTTGTTGAACGAA CCGAACGAAAAACTG
Rbg40.1F_mouse_combo_r4	CACCATCCCCGTCGATCTTCYCAGTAAACTGCCAAGACTGCATGGTGGTTWKGAT CMTCRMCAGTTTTTCGTTCGGTTCGT
Rbg40.1F_mouse_combo_f5	GAAGATCGACGGGGATGGTGCGCAGGAACTCGCGAAAGAAGTTGAAGAACTG
Rbg40.1F_mouse_combo_r6	CCTCAACGTCGATACCACGTWCTTCGTGTTCCTGTTSCAGTTCTTCAACTTCTTT CG
Rbg40.1F_mouse_combo_f7	ACGTGGTATCGACGTTGAGGACTACGCGTCTAACCTGAAAGTTATCCTGCTG
Rbg40.1F_mouse_combo_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGAGCCAGSTSCAGCAGGATAACTTTC AGGT
Rbg40.1F_human_combo_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCTACCAAAAAAACCCAGCTGCTG
Rbg40.1F_human_combo_r2	GCGCGTCCAGGAGCGCGTGTTSSAWCAGCAGCTGGGTTTTTTTGG
Rbg40.1F_human_combo_f3	ACACGCGCTCCTGGACGCGCWGVACATGARSARGRASTTGCCGGAACCGAACGAA AAACT
Rbg40.1F_human_combo_r4	TAAAGATCCAAGACTGCATGKYGGTGATGATACGGYTCAGTTTTTCGTTCGGTTC CGG
Rbg40.1F_human_combo_f5	CATGCAGTCTTGGATCTTTACTGRGAAGATCGACGGGGATGGTGCGCAGSASCTC GCGAAAGAAGTTGAAGA
Rbg40.1F_human_combo_r6	ACCACGTTTTTCGTGTTCCTGTTCCAGTTCTTCAACTTCTTTCGCGAG
Rbg40.1F_human_combo_f7	AGGAACACGAAAAACGTGGTSRGGASGTTGAGGACKASGCGTCTAACCTGAAAGT TATCCTGCT
Rbg40.1F_human_combo_r8	AAGCTTTTGTTCGGATCCGCCTCCCTCGAGAGCCAGCTCCAGCAGGATAACTTTC AGGTT
Rbg40.1F_human_combo_f1	AGCGGAGGCGGAGGGTCGGCTAGCCATATGTCTACCAAAAAAACCCAGCTGCTG

Table S3-2: Oligonucleotides used in the assembly of mouse and human combinatorial libraries based on Rbg40.1F. All sequences are written in 5'→ 3' orientation.

Mouse IL-2Rβ with basic leucine zipper
AVKNCSHLECFYNSRANVSCMWSHEEALNVTTCHVHAKSNLRHWNKTCELTLVRQASWAC NLILGSFPESQSLTSVDLLDINVVCWEEKGWRRVKTCDFHPFDNLRLVAPHSLQVLHIDT QRCNISWKVSQVSHYIEPYLEFEARRRLLGHSWEDASVLSLKQRQQWLFLEMLIPSTSYE VQVRVKAQRNNTGTWSPWSQPLTFRTRPADPMKEISRGGLEVLFQGPEFGGSTTAPSAQL KKKLQALKKKNAQLKWKLQALKKKLAQHHHHHH
Mouse γ_c with acidic leucine zipper
PLPEVQCFVFNIEYMNCTWNSSSEPQATNLTLHYRYKVSDNNTFQECSHYLFSKEITSGC QIQKEDIQLYQTFVVQLQDPQKPQRRAVQKLNLQNLVIPRAPENLTLSNLSESQLELRWK SRHIKERCLQYLVQYRSNRDRSWTELIVNHEPRFSLPSVDELKRYTFRVRSRYNPICGSS QQWSKWSQPVHWGSHTVEENSRGGLEVLFQGPEFGGSTTAPSAQLEKELQALEKENAQLE WELQALEKELAQGLNDIFEAQKIEWHEHHHHHH

Table S3-3: b-mIL2Rβγ heterodimer construct was formed by co-expressing IL-2Rβ ectodomain (residues 1-215) and γ_c ectodomain (residues 34-233) onto complementary leucine zipper base/acid pairs, with the latter containing a C-terminal BAP tag to facilitate biotinylation.

Primer	Templates	Sequence
5'_40F.M1,3	Rbg40F.M1, Rbg40F.M3	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCGAAA AAACCC
5'_40F.M2	Rbg40F.M2	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAA ATACCCAGC
3'_40F.M1	Rbg40F.M1	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGCCAGGTGC
3'_40F.M2,3	Rbg40F.M2, Rbg40F.M3	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGCCAGCTGC

5'_Rbg40.1	Rbg40F.H1 to Rbg40F.H6	AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGTCTACCAAAA AAACCCAG
3'_Rbg40.1	Rbg40F.H1 to Rbg40F.H6	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGAGCCAGCTCC AG

Table S3-4: Primers used to subclone genes encoding third-generation binders into pET29b. All sequences are written in 5'→ 3' orientation.

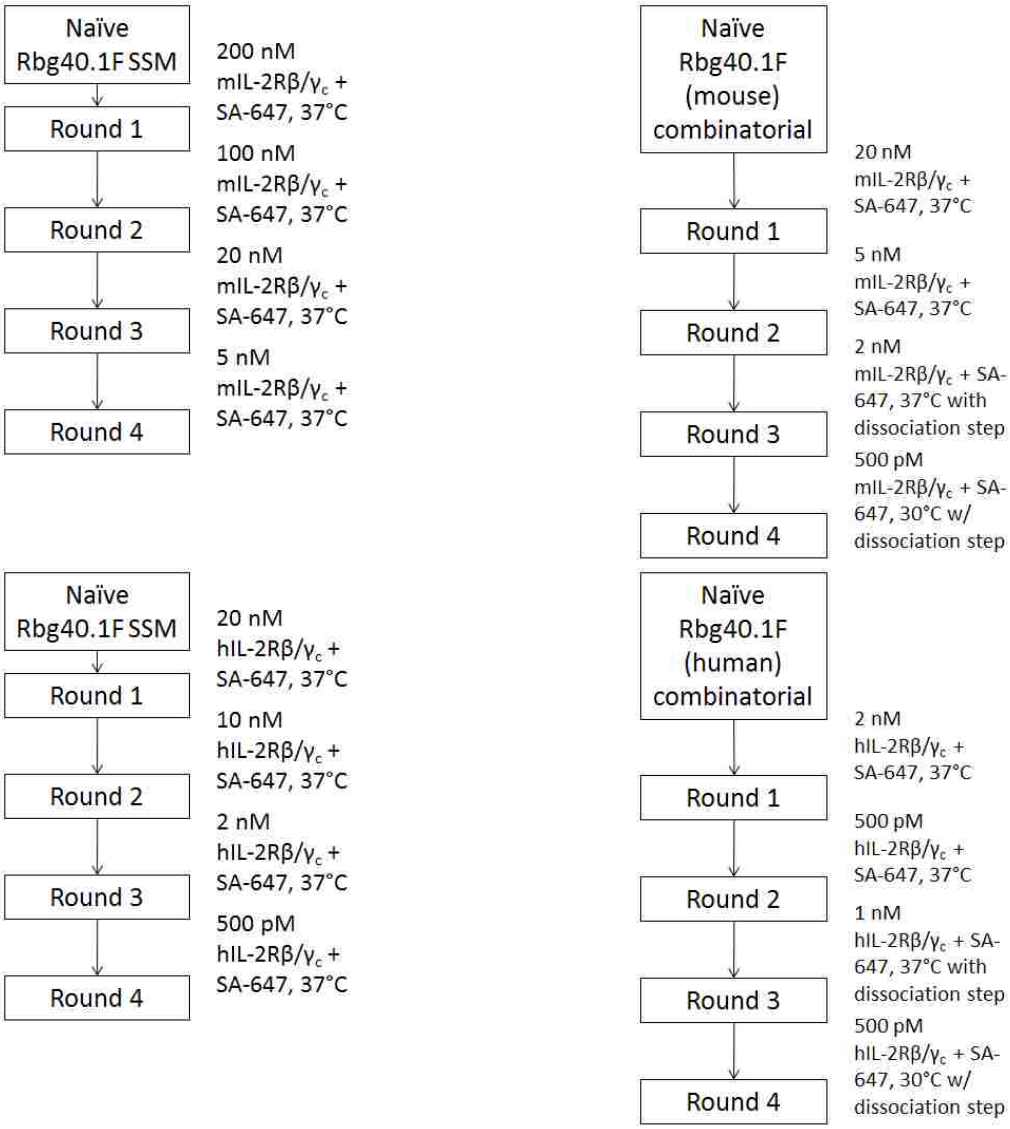


Figure S3-1: Selection schemes for Rbg40.1F SSM and combinatorial libraries against mIL-2Rβ/γ_c (top) and hIL-2Rβ/γ_c.

References

1. Regan, L. & DeGrado, W. Characterization of a helical protein designed from first principles. *Science (80-.).* **241,** 976–978 (1988).

2. Richardson, J. & Richardson, D. The *de novo* design of protein structures. *Trends Biochem. Sci.* **14,** 304–309 (1989).

3. Kaplan, J. & DeGrado, W. F. De novo design of catalytic proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 11566–11570 (2004).

4. Marvin, J. & Hellinga, H. Conversion of a maltose receptor into a zinc biosensor by computational design. *Proc. Natl. Acad. Sci.* **98,** 4955–4960 (2001).

5. Benson, D., Wisz, M. & Hellinga, H. Rational design of nascent metalloenzymes. *Proc. Natl. Acad. Sci.* **97,** 6292–6297 (2000).

6. Pinto, A., Hellinga, H. & Caradonna, J. Construction of a catalytically active iron superoxide dismutase by rational protein design. *Proc. Natl. Acad. Sci.* **94,** 5562–5567 (1997).

7. Dahiyat, B. & Mayo, S. De novo protein design: fully automated sequence selection. *Science (80-.).* **278,** 82–87 (1997).

8. Xiang, Z. & Honig, B. Extending the accuracy limits of prediction for side-chain conformations. *J. Mol. Biol.* **311,** 421–430 (2001).

9. Brooks, B., Bruccoleri, R., Olafson, B., Swaminathan, S. & Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4,** 187–217 (1983).

10. Liu, S. *et al.* Nonnatural protein--protein interaction-pair design by key residues grafting. *Proc. Natl. Acad. Sci.* **104,** 5330–5335 (2007).

11. Leaver-Fay, A. *et al.* ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzym.* **487,** 545–574 (2011).

12. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77,** 363–382 (2008).

13. Kuhlman, B. *et al.* Design of a novel globular protein fold with atomic-level accuracy. *Science (80-.).* **302,** 1364–1368 (2003).

14. Koga, N. *et al.* Principles for designing ideal protein structures. *Nature* **491,** 222–227 (2012).

15. Huang, P.-S. *et al.* High thermodynamic stability of parametrically designed helical bundles. *Science* **346,** 481–485 (2014).

16. Parmeggiani, F. *et al.* A general computational approach for repeat protein design. *J. Mol. Biol.* **427,** 563–575 (2015).

17. Park, K. *et al.* Control of repeat-protein curvature by computational protein design. *Nat. Struct. Mol. Biol.* **22,** 167–174 (2015).

18. King, N. *et al.* Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science (80-.).* **336,** 1171–1174 (2012).

19. King, N. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510,** 103–108 (2014).

20. Ashworth, J. *et al.* Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441,** 656–659 (2006).

21. Thyme, S. *et al.* Exploitation of binding energy for catalysis and design. *Nature* **461,** 1300–1304 (2009).

22. Tinberg, C. *et al.* Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501,** 212–216 (2013).

23. Mills, J. *et al.* Computational design of an unnatural amino acid dependent metalloprotein with atomic level accuracy. *J. Am. Chem. Soc.* **135,** 13393–13399 (2013).

24. Procko, E. *et al.* Computational design of a protein-based enzyme inhibitor. *J. Mol. Biol.* **425,** 3563–3575 (2013).

25. Röthlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453,** 190–195 (2008).

26. Jiang, L. *et al.* De novo computational design of retro-aldol enzymes. *Science (80-.).* **319,** 1387–1391 (2008).

27. Siegel, J. B. *et al.* Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science (80-.).* **329,** 309–313 (2010).

28. Gordon, S. *et al.* Computational design of an α-gliadin peptidase. *J. Am. Chem. Soc.* **134,** 20513–20520 (2012).

29. Rajagopalan, S. *et al.* Design of activated serine-containing catalytic triads with atomic-level accuracy. *Nat. Chem. Biol.* **10,** 386–391 (2014).

30. Fleishman, S. J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332,** 816–821 (2011).

31. Whitehead, T. A. *et al.* Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing. *Nat. Biotechnol.* **30,** 543–548 (2012).

32. Strauch, E.-M., Fleishman, S. J. & Baker, D. Computational design of a pH-sensitive IgG binding protein. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 675–680 (2014).

33. Procko, E. *et al.* A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells. *Cell* **157,** 1644–1656 (2014).

34. Fleishman, S. J. *et al.* Hotspot-Centric *De Novo* Design of Protein Binders. *J. Mol. Biol.* **413,** 1047–1062 (2011).

35. Correia, B. E. *et al.* Computational Design of Epitope-Scaffolds Allows Induction of Antibodies Specific for a Poorly Immunogenic HIV Vaccine Epitope. *Structure* **18,** 1116–1126 (2010).

36. Correia, B. *et al.* Proof of principle for epitope-focused vaccine design. *Nature* **507,** 201–206 (2014).

37. Leonard, W. Cytokines and immunodeficiency diseases. *Nat. Rev. Immunol.* **1,** 200–208 (2001).

38. Liao, W., Lin, J.-X. X. & Leonard, W. J. Interleukin-2 at the Crossroads of Effector Responses, Tolerance, and Immunotherapy. *Immunity* **38,** 13–25 (2013).

39. Abrams, D. *et al.* Interleukin-2 therapy in patients with HIV infection. *N. Engl. J. Med.* **361,** 1548–1559 (2009).

40. Tang, Q. *et al.* In vitro--expanded antigen-specific regulatory T cells suppress autoimmune diabetes. *J. Exp. Med.* **199,** 1455–1465 (2004).

41. Liu, R. *et al.* Expansion of regulatory T cells via IL-2/anti-IL-2 mAb complexes suppresses experimental myasthenia. *Eur. J. Immunol.* **40,** 1577–1589 (2010).

42. Chen, C. *et al.* IL-2 Simultaneously Expands Foxp3+ T Regulatory and T Effector Cells and Confers Resistance to Severe Tuberculosis (TB): Implicative Treg--T Effector Cooperation in Immunity to TB. *J. Immunol.* **188,** 4278–4288 (2012).

43. Harrison, C., Verstovsek, S., McMullin, M. & Mesa, R. Janus kinase Inhibition and its effect upon the therapeutic landscape for myelofibrosis: from palliation to cure? *Br. J. Haematol.* **157,** 426–437 (2012).

44. Kontzias, A., Kotlyar, A., Laurence, A., Changelian, P. & O'Shea, J. Jakinibs: a new class of kinase inhibitors in cancer and autoimmune disease. *Curr. Opin. Pharmacol.* (2012).

45. Levin, A. M. *et al.* Exploiting a natural conformational switch to engineer an interleukin-2 'superkine'. *Nature* **484,** 529–533 (2012).

46. Mitra, S. *et al.* Interleukin-2 Activity Can Be Fine Tuned with Engineered Receptor Signaling Clamps. *Immunity* **42,** 826–838 (2015).

47. Azoitei, M. *et al.* Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold. *Science (80-.).* **334,** 373–376 (2011).

48. Wang, X., Rickert, M. & Garcia, K. C. Structure of the Quaternary Complex of Interleukin-2 with Its α, ß, and γc Receptors. *Science (80-.).* **310,** 1159–1163 (2005).

49. Zhang, Y. & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33,** 2302–9 (2005).

50. Obita, T. *et al.* Structural basis for selective recognition of ESCRT-III by the AAA ATPase Vps4. *Nature* **449,** 735–739 (2007).

51. Spangler, J. Personal communication.

52. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One* **6,** e20161 (2011).

53. Adriano-Silva, D. Personal Communication.

54. Gront, D., Kulp, D. W., Vernon, R. M., Strauss, C. E. M. & Baker, D. Generalized fragment picking in Rosetta: design, protocols and applications. *PLoS One* **6,** e23294 (2011).

55. Hoover, D. M. & Lubkowski, J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30,** e43 (2002).

56. Chao, G. *et al.* Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1,** 755–768 (2006).

57. Benatuil, L., Perez, J. M., Belk, J. & Hsieh, C.-M. An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng. Des. Sel.* **23,** 155–9 (2010).

58. Gibson, D. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

59. Bloom, J., Labthavikul, S., Otey, C. & Arnold, F. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci.* **103,** 5869–5874 (2006).

60. Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.3r1. (2010).

61. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30,** 614–620 (2014).

62. Thomsen, M. C. F. & Nielsen, M. Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res.* **40,** W281–7 (2012).

63. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16,** 276–277 (2000).