Rational Engineering of Erythropoietin for Smarter Protein Therapeutics: Structure–Function Relationships and Molecular Geometry

Citation

Lee, Jung Min. 2020. Rational Engineering of Erythropoietin for Smarter Protein Therapeutics: Structure–Function Relationships and Molecular Geometry. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Permanent link

https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37368913

Terms of Use

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story.

Accessibility

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Chemistry & Chemical Biology

have examined a dissertation entitled:

Rational Engineering of Erythropoietin for Smarter Protein Therapeutics: Structure–Function Relationships and Molecular Geometry
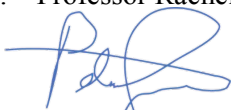
presented by: Jung Min Lee

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature _____

Typed name: Professor Pamela Silver

Signature _____

Typed name: Professor David Liu

Signature _____

Typed name: Professor Rachelle Gaudet

Signature _____

Typed name: Professor Peter Sorger

Date: 29 July 2020

Rational Engineering of Erythropoietin for Smarter Protein Therapeutics: Structure–Function Relationships and Molecular Geometry

A dissertation presented

by

Jung Min Lee

to

The Department of Chemistry and Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemistry

Harvard University

Cambridge, Massachusetts

July 2020

Rational Engineering of Erythropoietin for Smarter Protein Therapeutics: Structure–Function Relationships and Molecular Geometry

Abstract

Protein therapeutics exploit the biology of naturally occurring proteins to deliver desired pharmacological activities to the patient's body, and their use has increased dramatically over the past three decades. They have the advantages of reducing off-target effects, thanks to their higher structural diversity, and more specific target protein interactions compared to small molecule drugs. Despite their similarity to natural molecules and low off-target interactions, protein therapeutics can cause undesired effects by signaling on wrong tissues, triggering pleiotropic actions, or activating immune responses. This dissertation aims to demonstrate strategies to rationally engineer proteins to maximize drug efficacy while minimizing such side effects, using erythropoietin (EPO) as a model protein. Chapter 2 presents a bifunctional AND-gate ligand that directs EPO activity to target cells and away from other cells that also bear EPO receptors. Tissue-targeted activation of EPO receptors was achieved by modulating receptor–ligand binding strength and quantitatively analyzing the molecular geometry of protein domains on the target cell surface. Chapter 3 describes EPO fusion proteins that are designed to exhibit a subset of the pleiotropic activities of EPO based on its structure–function relationships in different signaling contexts, such that only desired effects can be selectively exploited for a given medical condition. Through this method, EPO fusion proteins with one or both of erythropoietic and tissue-protective effects were constructed. Chapter 4 delineates protein resurfacing approaches to reduce immune responses to EPO and prevent adverse consequences mediated by neutralizing anti-drug antibodies that are cross-reactive with its endogenous counterpart. Taken together, the present work shows that structural and spatial information about proteins can drive smarter therapeutic designs, as well as contributes to expanding the methods of protein engineering in drug development and construction of protein-based synthetic-biological systems.

Table of Contents

Acknowledgements

First and foremost, I would like to thank Professor Pamela Silver for her support, guidance, and mentorship throughout the last five years. My graduate school experience was not only educational but also stimulating and inspiring. I also thank Dr. Jeff Way for showing how fascinating the field of protein engineering can be and sharing his knowledge and wisdom as a protein whisperer.

I feel very lucky to have been surrounded by brilliant colleagues and friends in the Silver Lab. They have created such friendly, fun, and constructive environment. Special thanks to Katherine Redfield and Yukiye Koide for being passionate research partners, fun coffee-mates, and most importantly, amazing friends full of creative ideas for science and other funzies. Thank you to Nathalie Gruber for being such a hardworking mentee and research partner as well as a warmhearted, caring friend. To Juline Savigny for being a true friend who dragged me out of my comfort zone and pushed me to live a healthier life by being the best jogging buddy and trainer I could ask for. Thank you to Devin Burrill for her mentorship during my rotation, and to Marika Ziesack for being an awesome baymate and helping me transition to the Silver Lab and the Wyss. Thank you to Lauren Jiang and Kasia Kready for their valuable comments on my thesis.

I could not have accomplished my PhD work without my valuable collaborators. Thank you to Andyna Vernet for performing the animal work for my projects. Thank you to Professor Ionita Ghiran and Shulin Lu for teaching me red blood cell biology and relevant experiments, and letting me use their equipment.

Many thanks to friends outside the lab as well. Especially to Minjoo Jang, Sangjung Kim, Linda Kim, Sally Park, Jeannie Chung, Ayoung Keum, and Jongwoo Nam for helping me maintain life outside the lab. I also thank Professor Jiyong Hong for his continuous mentorship, guidance, and motivation regardless of distance.

Lastly, my deepest gratitude goes to my dear family and fiancé. Special thanks to my parents who have taught me to be adventurous and have always supported my decisions and passion, even if that meant sending me to the other side of the globe. To my brother who has always been my best friend and the sweetest brother one could ask for. To my fiancée who has always stood by me and filled my life with constant laughter and joy.

Chapter 1

Introduction

1.1 Protein Therapeutics: Strengths and Limitations

Since recombinant DNA technology was developed in the 1970s, the creation and commercialization of protein therapeutics has grown very rapidly. The first protein drug, recombinant human insulin, was approved by the U.S. Food and Drug Administration (FDA) in 1982. These drugs have become a mainstay of modern therapy. As of 2018, seven out of the top ten bestselling drugs were proteins [1].

Protein therapeutics have several advantages over most other types of drugs. Protein–protein interactions are potent and highly specific. They can be further engineered to mimic or improve natural protein interactions and fine-tune target specificity. Our knowledge of evolution and structure–function relationships of proteins enables rational design of protein therapeutics with desired properties. Unlike the majority of small molecule drugs, proteins are naturally found in our body, making it easier to predict their behaviors, such as pharmacokinetics and toxicity. Compared to small molecules and cell-based therapeutics, the manufacturing process for proteins is simpler, more controllable, and more predictable.

One class of protein therapeutics is cytokines and hormones. Both are signaling proteins that can modulate immune responses or regulate cell maturation and growth. Cytokines normally act locally, such as at the site of a wound or in a specific tissue, while hormones act throughout the entire body. When given as drugs, they can provide a replacement for endogenous proteins that a patient lacks or makes incorrectly due to a genetic disorder. They can also supplement or augment

cytokine/hormone activity in a patient who produces an insufficient amount of endogenous protein or needs more than standard levels. For example, insulin (Humulin®) and erythropoietin (EPO, Epogen®) are used to treat diabetes and anemia, respectively, by providing a replacement for or supplementing the endogenous protein. Interleukin-2 (IL-2, Proleukin®), tumor necrosis factor alpha (TNFα, Beromun®), and interferon alpha (IFNα, Intron A®) augment normal immune responses to treat cancer.

Another class of protein therapeutics is monoclonal antibodies. Monoclonal antibodies are large proteins produced by white blood cells to neutralize or attack specific targets as part of our adaptive immune system. Borrowing from these natural therapeutics, monoclonal antibody-based drugs can block undesired native signaling or trigger immune responses in patients. Adalimumab (Humira®), for instance, is an anti-TNFα antibody that mitigates excessive immune responses in rheumatoid arthritis by inhibiting the binding of this inflammatory signaling molecule to its receptor. Rituximab (Rituxan®) is an antibody to CD20, a cell surface marker on B cells, and mediates killing of B cells to treat B cell lymphomas, autoimmune diseases, and rheumatoid arthritis. Monoclonal antibodies have the advantages of their large size conferring long half-life in the body (whereas proteins with a molecular weight of less than about 50,000 Da are filtered by the kidney into the urine) [2, 3] as well as highly specific target affinity making them optimal for neutralizing undesirable antigens while avoiding most off-target side effects.

In spite of these advantages, protein drugs also often have particular downsides. The administration of 'natural' cytokines as drugs and their distribution throughout the body present an 'unnatural' situation, as naturally formed cytokines act locally, and can cause serious side effects. Most cytokines are associated with severe dose-limiting toxicities that come from on-target binding to non-target cell types. For example, IL-2 causes fever and a general upregulation of the immune system, and IFNα causes 'flu-like symptoms' that can be difficult to bear when this drug is given for a year (as is typical in treatment for hepatitis viruses) [4]. EPO's anemia-relieving red blood cell (RBC) replenishment is paired with blood clotting and vessel growth [5]. These last two effects are adaptive features in the natural setting of wound repair, but become extremely risky in generic anemia treatment. When EPO is given for years to patients with kidney failure, these effects cause an increase in heart attacks, strokes, and deep vein thromboses [6, 7, 8]. Cancer patients receiving EPO may get relief from their anemia-induced exhaustion, but have increased

blood supply to tumors [5, 9]. The challenge for synthetic biologists is the ability to manipulate the properties of these drugs through careful, biologically inspired design so that side effects can be minimized.

1.2 Protein Engineering Strategies to Reduce Side Effects

Thanks to recombinant DNA technology and protein engineering, researchers have learned to mutate, truncate, or attach protein parts to enhance the clinical potential of protein therapeutics. Etanercept (Enbrel®) was the first fusion protein drug to be approved by the FDA in 1998. It is composed of a fragment of TNF receptor and the dimeric Fc domain of an antibody, putting two receptor domains into a binding configuration so that they can sequester TNF, a profoundly inflammatory cytokine. Etanercept also has a longer plasma half-life compared to the unfused receptor fragment because of its higher molecular weight. Whole antibody–cytokine fusion proteins do even more than extend half-life — they are composed of the antigen-binding domain of an antibody connected via a linker to the active cytokine domain, such that the antibody fragment leads the cytokine signaling specifically to the desired target cell types. In one preclinical study, an antibody–IL12 fusion protein directed to the tumor microenvironment showed potent activity in various mouse cancer models at a dose that was more than 20-fold lower than the unfused cytokine alone [10]. However, the cytokine domain is not prohibited from binding to its receptors on non-target cell types. Clinical studies suggest that side effects are as dose-dependent as desired effects, and result from cytokine action on non-target tissues [10]. Attempts to increase half-life by definition cause longer exposure times, and improved tumor penetration may require higher doses, leading to more opportunity for undesired side-effects.

A "chimeric activator" is a type of antibody fusion protein that incorporates a mechanism to prevent the activating cytokine/hormone domain from binding to non-target cell receptors [11]. This is achieved by tuning the receptor binding of the activating domain via mutation so that signaling is reduced. The binding becomes too weak to activate its receptor in most tissues, and activity is rescued only when the cytokine is highly concentrated at its desired target location by the antibody element binding to a second protein on the target cells. This anchoring thereby shifts the equilibrium towards the receptor-bound state. This design allows for cell-targeted

3

delivery, target cell-specific activity, and removal of side effects mediated by on-target signaling on non-target cells. The general design principle of chimeric activators will be used to engineer tissue-targeted forms of EPO with reduced side effects. This work will be further discussed in Chapter 2, and will lay the foundation for the work described in Chapters 3 and 4.

1.3   Biologcial Roles and Structure of EPO

EPO was first discovered as the protein hormone responsible for the production of RBCs during studies on erythropoiesis and anemia in the early–mid 1900s, and hence the name 'erythropoietin' [12]. Once synthesized primarily in the kidneys, EPO travels to the bone marrow, and provides essential signals for the late RBC precursor cells to survive and commit completely to the erythroid lineage [12]. EPO is also involved in a set of other biological functions, such as blood clotting, angiogenesis, and tissue protection, all of which are adaptive features in the context of wound healing [5].

EPO is composed of 166 amino acids that form a four-helical bundle connected via flexible loops. It has one *O*-linked and three *N*-linked glycosylation sites that are decorated with oligosaccharides after translation and contribute to improved serum half-life and in vivo bioactivity of EPO [13]. The X-ray crystal structure of EPO in complex with an EPO receptor (EPO-R) homodimer has been solved (PDB ID: 1CN4 and 1EER), showing that EPO has two asymmetric receptor-binding sites that each bind to EPO-R via the same ligand-binding site with different affinities [14, 15]. The dissociation constant (K_D) of site 1, or the strong side, is about 1 nM, whereas the K_D of site 2, or the weak side, is about 1 µM [16]. The initial binding of EPO to EPO-R is likely driven by the strong side, and the complete signaling complex forms when another EPO-R monomer binds to the weak side and locks the conformation into the one that is amenable to JAK2 dimerization and downstream signaling in the cytoplasm. However, whether EPO-R dimerizes prior to EPO binding or EPO binding induces receptor dimerization is still controversial [17, 18, 19]. Upon binding to EPO-R, EPO signals primarily through the JAK2–STAT5 signaling pathway that leads to upregulation of anti-apoptotic factors, such as Bcl-X_L, as well as through other pro-proliferative and anti-apoptotic pathways involving STAT1/3, PI3K, and MAPK [20].

Due to its RBC-proliferative effects, EPO naturally lends itself to commercialization and

clinical application for the treatment of anemia. The first recombinant EPO was successfully synthesized in 1985, and synthetic EPO was first used to successfully treat anemia in 1987 [21, 22]. The FDA approved recombinant EPO for use in anemias associated with chronic renal diseases in 1989, and subsequently, in those associated with cancer in 1993 [23, 24]. Since then, several versions of recombinant EPO and its derivatives, such as Epogen®, Procrit®, Aranesp®, and Mircera®, have been commercialized and are widely used for the treatment of anemia in renal failure patients and cancer patients undergoing chemotherapy. However, in a set of clinical trials performed in the mid-2000s, such erythropoiesis-stimulating agents (ESAs) were shown to accompany increased risk of death, stroke, heart attack, deep vein thrombosis, and recurrence of cancer [6, 7, 8, 9]. Consequently, the FDA added a black box warning for these life-threatening adverse effects, limited the recommended dose of ESAs for renal patients, and prohibited the use of ESAs in cancer patients who are likely to achieve five-year survival [24].

1.4 Chapter Summary

The goal of this dissertation is to rationally design protein therapeutics by understanding the structure–function relationships and molecular geometry of protein domains. This dissertation is focused on engineering EPO to create safe drugs for various medical conditions. In Chapter 2, tissue-targeted forms of EPO are designed and optimized to stimulate RBC production without causing pro-thrombotic side effects. This chapter demonstrates the importance of understanding the molecular geometry of protein domains in designing multi-component fusion proteins. In Chapter 3, tissue-protective forms of EPO are designed to address diverse medical needs, such as high altitude-related illnesses, organ transplantation, and neurodegenerative diseases. In doing so, structural alignments were performed to generate structural models for a tissue-protective signaling complex of EPO, and to enhance our understanding of its structure–function relationships in the context of tissue protection. Chapter 4 discusses protein engineering strategies to reduce the immunogenicity of recombinant EPO drugs and to prevent anti-drug antibodies from cross-reacting with endogenous protein in the case of an immune response. The concluding chapter closes this dissertation by highlighting the promise of protein engineering in the development of smarter therapeutics as well as in a broader scope of mammalian synthetic biology.

Chapter 2

Rational Design of a Bifunctional AND-Gate Ligand to Modulate Cell–Cell Interactions

2.1 Abstract

Protein "AND-gate" systems, in which a ligand acts only on cells with two different receptors, direct signaling activity to a particular cell type and avoid action on other cells. In a bifunctional AND-gate protein, the molecular geometry of the protein domains is crucial, as fusion proteins composed of two or more domains have the potential to crosslink cells, whether desired or not. In this chapter, I constructed a tissue-targeted erythropoietin (EPO) that stimulates red blood cell (RBC) production without triggering thrombosis. The EPO was directed to RBC precursors by fusion to an anti-glycophorin A (GPA) antibody V region. Many such constructs activated EPO receptors in vitro, and stimulated RBC but not platelet production in mice. However, they also caused adhesion between RBCs and EPO receptor-bearing cells in vitro, and enhanced thrombosis in mice. I found that the linker element of the fusion protein, the position of the antibody binding epitope, and the target cell surface proteins synergistically contribute to cell–cell interaction mediated by the fusion protein. Based on the protein structural model of the RBC surface, I rationally designed a form of an anti-GPA/EPO fusion protein that does not mediate

cell–cell adhesion in vitro or enhance thrombosis in mice. This work shows how meso-scale geometry can inform design of synthetic-biological systems.

2.2 Introduction

2.2.1 Multi-Domain Fusion Proteins in Synthetic-Biological Systems

Engineered proteins in a synthetic-biological system often comprise two or more domains, such that several desired features can be implemented in a single engineered protein. Such designs enable us to harness natural protein functions more effectively by directing the protein to desired locations and by enhancing the desired activity via synergy between the spatially connected proteins. The most well-studied examples include chimeric antigen receptor T (CAR-T) cells, bispecific T cell engagers (BiTEs), and antibody–cytokine fusion proteins (immunocytokines). CAR-T cells specifically kill tumor cells that present target antigens on the cell surface, and BiTEs physically bridge T cells and target cells to facilitate target cell killing (Figure 2.1A) [25, 26]. For immunocytokines, fusion to an antibody fragment can direct cytokine activity to target cells, increase serum half-life, reduce immunogenicity, or load cytokines on red blood cells (RBCs) like a drug delivery platform (Figure 2.1B,C) [27, 28, 29, 30].
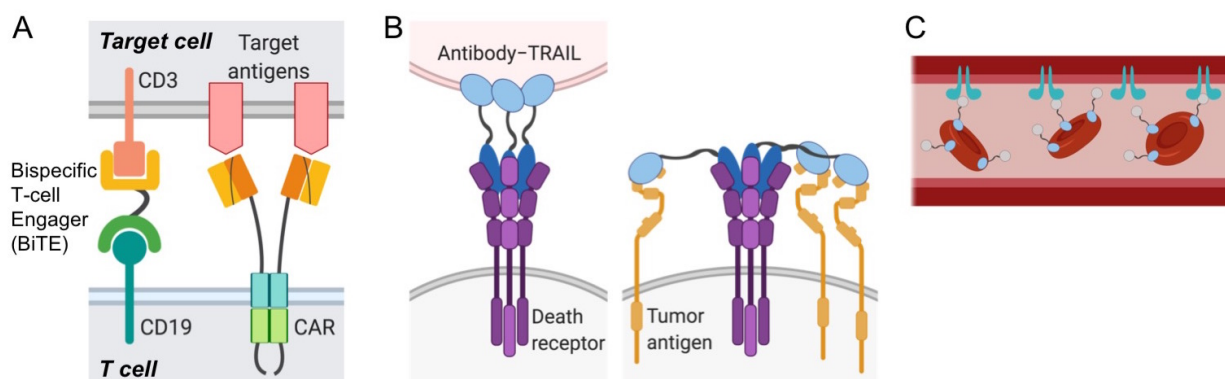
Figure 2.1: *Multi-domain fusion proteins in synthetic-biological systems. (A) Fusion proteins, such as BiTEs and CARs, take advantage of cell–cell crosslinking to induce target cell-specific killing. (B) Immunocytokines, such as antibody–TRAIL fusion proteins, can benefit from leveraging cell–cell crosslinking. (C) Immunocytokines that hijack on RBCs for effective drug delivery can show unwanted side effects caused by cell–cell crosslinking or signaling that occurs in trans in the blood vessels. Images were created with BioRender.com.*

Multi-domain fusion proteins are capable of binding to two or more proteins at the same

time, and thus, have the potential to crosslink cells, whether desired or not. For CAR-T cells and BiTEs, their fusion proteins are intended to connect target cells to be killed with T cells, such that the two cells can form an immunological synapse essential for effective T cell-mediated killing (Figure 2.1A). Crosslinking of two or more cells can be beneficial for some immunocytokines as well. For example, TRAIL triggers cell death more potently when in membrane-bound form than when in soluble form, as it is then able to form tight cell–cell contacts that lead to high-order oligomeric receptor complex formation [31, 32]. When TRAIL is fused to an antibody fragment targeted to the RBC surface, it acts on its target cells in trans, mimicking the action of the membrane-bound form and outperforming another TRAIL fusion protein designed to act in cis on the target cell surface (Figure 2.1B). In contrast to these examples, cell–cell crosslinking may be undesired and lead to adverse effects in some scenarios. For instance, immunocytokines that use RBCs as a delivery platform should consider the possibility of unintended receptor binding and signaling that could occur in trans between RBCs and nearby cells in blood vessels (Figure 2.1C). Therefore, understanding the structural determinants for engineered cell–cell interaction is important for designing the next-generation CAR-T cells, BiTEs, and other fusion proteins.

2.2.2   Chimeric Activator for Target Cell-Specific Drug Action

Naturally, hormones and cytokines act on receptors on diverse cells to achieve coordinated biological responses. When these hormones and cytokines are given as drugs, some activities may be desired while others may constitute drug side effects. For example, cortisol acts on the glucocorticoid receptor in immune cells and is anti-inflammatory, but also acts on other cells to increase blood sugar and induce bone resorption. Similarly, erythropoietin (EPO) stimulates RBC production but also increases thrombotic activity. These responses are adaptive in the context of hemorrhage, but increased thrombotic activity may underlie the higher frequency of heart attacks and strokes seen in EPO-treated kidney failure patients.

The Silver Lab previously designed a class of AND-gate proteins, "Chimeric Activators," in which a cytokine or hormone is mutated to reduce its activity and then attached to a targeting element that binds to another protein on a subset of receptor-bearing cells (Figure 2.2A–D) [11]. Chimeric activators are similar to other fusion proteins such as blinatumomab [25] or

immunocytokines [27]. Thus, they have the potentials to crosslink cells, and should be carefully designed to leverage or avoid such events.
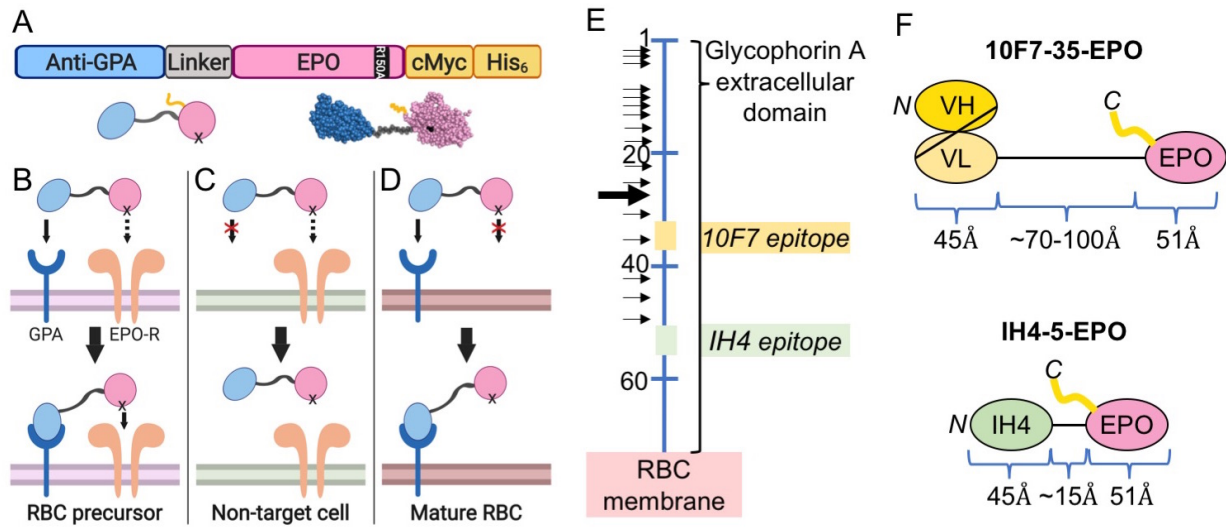
Figure 2.2: *(A) Protein sequence organization of a Targeted EPO chimeric activator. (B) The chimeric activator binds to a target cell via the targeting element and tethers the mutated activity element to the cell surface so that the mutated activity element binds to its receptor despite the mutation. (C) On a non-target cell that lacks the surface marker to which the targeting element binds, the mutated activity element has little effect. (D) Targeted EPO molecules also bind to mature RBCs via GPA, creating a sink for a Targeted EPO to extend its plasma half-life. EPO-R is not present on these cells. (E) Extracellular portion of GPA (blue line), showing amino acid positions (numbers), O-linked (small arrows) and N-linked (large arrow) glycosylation sites, and the positions of the 10F7 and IH4 epitopes. (F) The 10F7-35-EPO(R150A) and IH4-5-EPO(R150A) forms of Targeted EPO, showing sizes in angstroms (Å) based on typical scFv and nanobody sizes, an extended conformation of the Gly/Ser linkers, and solved structures of EPO. Some images were created with BioRender.com.*

Based on the chimeric activator design, an AND-gated EPO that activates EPO receptors (EPO-Rs) on RBC precursors but not on other cells was previously constructed [11, 33]. "Targeted EPO" proteins consist of an antibody element that binds to glycophorin A (GPA), EPO with a mutation that reduces its activity, and a connecting linker (Figure 2.2A) [11, 33]. GPA is abundant on RBCs (800,000 copies/cell) and late RBC precursors (50,000 copies/cell), but is not expressed on cells that control thrombosis and tumor angiogenesis, or on solid tumor cells [33, 34]. EPO-Rs on late RBC precursors mediate EPO-dependent maturation, but then disappear as cells mature [33, 35]. Targeted EPO was designed to bind to late RBC precursors and stimulate EPO-Rs on only those cells (Figure 2.2B). On non-target cells, it fails to activate EPO-Rs because the mutated EPO alone binds poorly to receptors (Figure 2.2C). Binding to GPA on mature RBCs is a desirable secondary outcome, as it extends the plasma half-life of the fusion protein (Figure 2.2D) [33].

The previously characterized form of Targeted EPO has a long linker between the antibody and EPO elements, and an anti-GPA antibody element that binds to a membrane-distal epitope. It was able to stimulate production of RBCs but not platelets, with the latter as an indicator of non-target cell effects. While this result indicated cell-type specificity of action, platelet number alone is not strongly correlated with thrombotic propensity, and the possibility remained that the fusion protein might still cause thrombosis despite its seeming cell-type specificity. As such I found that the initial fusion protein might have a thrombotic activity in vivo. The structure–function analysis presented here indicates that, depending on the geometry of Targeted EPO, the EPO moiety tethered to mature RBCs could interact with EPO-Rs on other cells and lead to cell–cell adhesion, but that the potential for such interaction could be eliminated by rational design. In vivo, a poorly designed fusion protein could bind to both RBCs and endothelial cells [36] or immune cells such as macrophages [37] and mediate undesired cell–cell adhesion or signaling [33]. The goal of the present work is to define structural and quantitative features of anti-GPA/EPO chimeric activators that will act specifically in cis on GPA-expressing cells, to avoid side effects that may arise from unwanted cell–cell interaction.

2.3 Results and Discussion

Previously the Silver Lab characterized a form of Targeted EPO consisting of V regions from the 10F7 antibody, a 35-amino acid linker, and EPO with the weakening mutation R150A; this version was termed 10F7-35-EPO(R150A) (Figure 2.2F). It was able to activate EPO-R in a target cell-specific manner in vitro and in mice [33]. To directly assess the impact of this molecule on thrombosis, I performed a 'bleeding time' assay, in which the tail of a treated mouse is transected and the time to cessation of bleeding is measured.

Contrary to expectation, 10F7-35-EPO(R150A) promoted thrombosis in the bleeding time assay (Figure A.1). I hypothesized two mechanisms by which this might occur. First, the antibody element could, by itself, promote blood clotting. Second, the RBC-bound Targeted EPO could crosslink RBCs with other cells bearing EPO-Rs, such as vascular endothelial cells or leukocytes, and thus slow blood flow. To determine whether either or both of these hypotheses were correct, I constructed a series of Targeted EPO variants, testing anti-GPA antibody elements (10F7, 1C3, R18

10

and IH4), fusion protein configurations, and linker lengths.

2.3.1 Targeted EPO Activates EPO-Rs in a GPA-Dependent Manner in Vitro and in Vivo

Diverse forms of Targeted EPO activate EPO-Rs in a GPA-dependent manner in vitro. TF-1 cells, which express both GPA and EPO-Rs [11], were treated with several forms of Targeted EPO, and their proliferation was measured 72 hr post-treatment (Figure 2.3A,B). Unfused EPO(R150A) showed about 100-fold less activity compared to wild-type EPO (EPO(WT)). When EPO(R150A) was fused to an anti-GPA antibody element, its activity was rescued by about 10- to 100-fold (Figure 2.3A,B). IH4 and 1C3 were the most potent antibody elements tested.

Protein (V region–linker–EPO)	N	Activity in vitro Log(EC$_{50}$) ± SD	EC$_{50}$ relative to epoetin alfa
Epoetin alfa (Amgen)	16	-10.239 ± 0.345	1
Darbepoetin (Amgen)	8	-8.989 ± 0.135	14.67
EPO(WT)	3	-10.378 ± 0.528	0.87
EPO(R150A)	7	-8.078 ± 0.233	130.86
EPO(K45D)	3	-6.572 ± 0.754	7298
10F7-35-EPO(R150A)	4	-8.947 ± 0.206	16.71
10F7-5-EPO(R150A)	8	-9.378 ± 0.754	13.69
10F7-EPO(R150A)-HEL	4	-9.142 ± 0.501	16.02
10F7-EPO(R150A)-LEH	4	-8.648 ± 0.643	54.76
1C3-35-EPO(R150A)	6	-10.517 ± 0.471	0.81
R18-17-EPO(R150A)	2	-9.415 ± 0.409	6.44
R18-5-EPO(R150A)	2	-9.495 ± 0.185	4.56
IH4-35-EPO(R150A)	16	-10.385 ± 0.756	1.45
IH4-5-EPO(R150A)	9	-10.410 ± 0.293	0.69
IH4-35-EPO(K45D)	5	-9.591 ± 0.364	4.70

Figure 2.3: *In vitro and in vivo activity of different forms of Targeted EPO. (A) Typical TF-1 cell proliferation assay comparing the stimulation of proliferation by "wild-type" EPO (EPO(WT); including C-terminal tags), the EPO(R150A) mutant protein, and the IH4-35-EPO(R150A) fusion protein. The effective concentration (EC$_{50}$) was calculated by a four-parameter fit. (B) Table summarizing the ability of several different forms of Targeted EPO and control proteins to promote proliferation of TF-1 cells in vitro. (C,D) An in vivo test of the ability of IH4-5-EPO(R150A) to specifically promote RBC production and not platelet production using transgenic mice that express human GPA on their RBCs [33]. The results indicate that IH4-5-EPO(R150A) and diverse other forms of Targeted EPO show enhanced activity of the EPO(R150A) element independently of the mode of attachment of EPO to the GPA-binding antibody element. Data represent mean ± S.E.M.*

Three configurations of Targeted EPO were tested to address whether restricting the

movement of EPO would affect its activity and potentially cause GPA-bound EPO to preferentially act in cis on targeted cells. The conventional scFv–EPO configuration is designed as (antibody heavy chain)–(light chain)–(EPO) ("HLE"). I modified it such that EPO is between the heavy and light chains; this is feasible because the N- and C-termini of EPO are nearby in the folded structure. The resulting configurations ("HEL" and "LEH") were expressed and tested. Both 10F7-EPO(R150A) and 1C3-EPO(R150A) stimulated proliferation of TF-1 cells largely independently of configuration, indicating that orienting EPO in one way does not significantly constrain its receptor binding in cis (Figure 2.3B and Table A.1).

The GPA-mediated enhancement of EPO(R150A) activity was independent of the linker length in the various fusion proteins. Changing the linker of 10F7-EPO(R150A), R18-EPO(R150A), R18-EPO(K45D), IH4-EPO(R150A), and IH4-EPO(K45D) from 5 to 35 amino acids (about 10–80 Å) in fusion proteins containing the same antibody and EPO elements showed similar EC_{50} values in cell-based assays (Figure 2.3B and Table A.1). Results with the rotationally-constrained and short-linker constructs are consistent with the idea that GPA is an intrinsically disordered, highly flexible protein, so that diverse fusion proteins bound to GPA can still allow binding of EPO with EPO-R.

When tested in vivo, various forms of Targeted EPO specifically stimulated production of RBCs but not platelets. The in vivo target cell specificity and efficacy of several Targeted EPO proteins were assessed in mice that are transgenic for human GPA, using darbepoetin, an extended half-life form of EPO, as a control to represent non-targeted EPO activity. Mice were injected intraperitoneally (i.p.) with vehicle, darbepoetin, or Targeted EPO, and their reticulocytes and reticulated platelets were measured to indicate activity on target cells and non-target cells, respectively. Mice treated with darbepoetin showed comparable increases in both new RBCs and new platelets (Figure 2.3C,D and Figure A.2) [33]. Mice treated with IH4-5-EPO(R150A) showed a similar increase in reticulocytes while the reticulated platelets were not significantly changed (Figure 2.3C,D). Of the forms of Targeted EPO tested, IH4-5-EPO(R150A) was the most potent in vivo (Figure 2.3C,D and Figure A.2) [33].

2.3.2 The Antibody Element IH4 Does Not Induce 'RBC Inflammation'

The anti-GPA antibody element in forms of Targeted EPO might induce an inflammatory phenotype that could enhance blood clotting or cause other undesired side effects as well as confound interpretation of experimental treatment with fusion proteins. Various anti-GPA antibodies and V region elements induce a constellation of pro-inflammatory phenotypes on RBCs, including increased membrane stiffness during shear stress, decreased membrane fluidity, rearrangement of RBC skeletal proteins, secretion of ATP, production of reactive oxygen species (ROS), and phosphorylation of Tyr8 of band 3. Divalent IgG-type antibodies against GPA induce at least some, if not all, of these phenotypes, but only a subset of monovalent Fabs, scFvs and VHHs induces this RBC response in an epitope-dependent manner (Figure 2.2E) [38, 39, 40, 41, 42, 43].
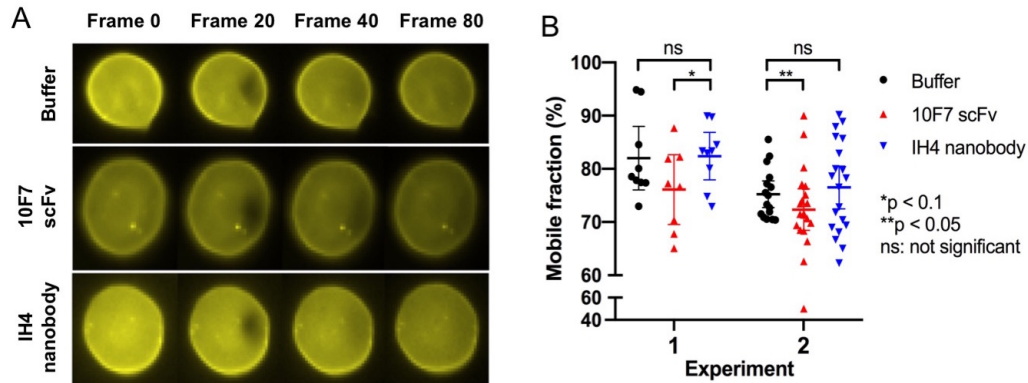
Figure 2.4: *Reduction in membrane fluidity by 10F7 but not IH4, as measured by FRAP. RBC membranes were stained with a lipid-soluble dye, a portion of the membrane was bleached with a laser, and diffusion of the dye into the bleached region was measured. (A) Serial snapshots of RBCs treated with antibody fragments during FRAP experiments. Each recording was made at eight frames per second for 10 s. (B) Mobile lipid fraction of eight or nine RBCs (Experiment 1) or 16–20 RBCs (Experiment 2) treated with buffer, 10F7, or IH4. When the data from these experiments were normalized to the buffer controls in the same experiment and then combined, the difference between 10F7 and buffer (p = 0.0648) or 10F7 and IH4 (p = 0.0463) was significant (Mann–Whitney test). Lines indicate means and 95% confidence intervals. The distribution of individual cell responses is typical for these assays [44].*

I used fluorescence recovery after photobleaching (FRAP) to assess whether various antibody elements induced an 'inflamed' state in RBCs. RBC membrane deformability and lipid mobility are some of the most frequently used readouts for such changes in RBCs upon treatment with stimuli, such as anti-GPA antibodies and complement. Piagnerelli *et al.*, 2007 showed that RBCs of septic patients had altered lipid bilayers and decreased membrane deformability [45]. Khoory *et al.*, 2016 illustrated that binding of particular anti-GPA antibodies, R10 or E4, induced

an increase in intracellular ROS, ATP release, RBC rigidity, a decrease in membrane deformability, and a decrease in the mobile lipid fraction as measured by FRAP [44]. Decreased membrane fluidity is correlated with RBC inflammation and may be a direct consequence of cell stiffness that results from cytoskeletal changes upon inflammatory signaling [42, 44]. For example, proteins such as GPA and band 3, whose transmembrane segments occupy a significant fraction of the RBC membrane, may become immobilized and thus slow the movement of the surrounding lipids.

The single-chain camelid element (VHH or nanobody) IH4 itself lacks stiffening/inflammatory activity when it is bound to GPA, making it a good candidate targeting element [42]. This element binds to the membrane-proximal epitope ($_{52}$YPPE$_{55}$) on human GPA [43]. I tested an optimized variant of IH4 (Figure A.3), and confirmed by FRAP that 10F7 induced lower membrane lipid mobility in RBCs while IH4 did not (Figure 2.4). In addition, IH4 appears to bind most tightly of the known anti-GPA V region elements (Table A.2). Therefore, IH4 was chosen as a candidate antibody element for a Targeted EPO that would not induce undesired signaling and would also limit potential interaction of EPO with EPO-Rs on other cells.

2.3.3 Molecular Geometry of GPA and Targeted EPO Determines Potential for Cell–Cell Crosslinking

In the design of Targeted EPO, the intention is that the antibody element and EPO bind in cis to receptors on the same cell surface (Figure 2.2B). However, a fusion protein comprising two binding elements could crosslink different cells that each bear a receptor for one of its components. To address this, I tested the accessibility of antibodies to the EPO element of Targeted EPO variants bound to RBCs as well as the ability of different Targeted EPOs to promote binding between RBCs and tumor cells expressing EPO-Rs.

When forms of Targeted EPO are bound to RBCs via GPA, the antibody accessibility of the EPO element decreases when the linker length is shortened and the epitope on GPA is closer to the cell membrane. I compared a series of proteins in which His$_6$-tagged EPO was fused to IH4 (epitope at $_{52}$YPPE$_{55}$) or 10F7 (epitope at about $_{34}$YAATP$_{38}$) and the linker length was varied. RBCs were incubated with the fusion protein, and then the accessibility of bound EPO was assessed via flow cytometry by binding of a phycoerythrin (PE)-conjugated anti-His$_6$ antibody. For both

Figure 2.5: *Targeted EPO fusion proteins with short linkers and membrane-proximal GPA binding epitopes reduce EPO accessibility and potential to bridge two surfaces. (A,B) Targeted EPOs containing either 10F7 or IH4 and varying linker lengths were incubated with RBCs, and stained with a PE-labeled anti-His$_6$ antibody. (A) Accessibility of the EPO element by anti-His$_6$-PE antibody, shown as the mode fluorescence intensity (PE). (B) Crosslinking of RBCs mediated by divalent anti-His$_6$ antibody, shown as forward scattering (particle size). A signal greater than the cutoff (dotted line) indicates adhesion of RBCs. (C,D) Petri dish was coated with Targeted EPOs containing either 10F7 or IH4 and either 35 or 5 amino acid linker, and was incubated with RBCs. (C) RBCs bound to the sections of petri dish that were coated with Targeted EPOs. (D) The number of RBCs bound to the coated petri dish, estimated by the area covered by RBCs. Data represent mean $\pm$ S.E.M.*

IH4 and 10F7 fusion proteins, PE signals progressively decreased as the linker was shortened (Figure 2.5A). In addition, crosslinking of RBCs by the divalent anti-His$_6$ antibody increased with linker length of the fusion protein, as evidenced in forward scattering signals by a shoulder to the right of the main peak (Figure 2.5B). Regardless of linker length, IH4 fusion proteins had lower accessibility and fewer RBC crosslinking events by the anti-His$_6$ antibody than 10F7 fusion proteins, suggesting that the binding epitope of the antibody element also affects the accessibility of EPO (Figure 2.5A,B).

When Targeted EPO fusion proteins are bound to a petri dish as a surrogate for the non-target cell surface, their ability to bind to GPA on RBCs also decreases with the linker length and the distance between the epitope on GPA and the RBC membrane. Parts of a petri dish were coated with forms of Targeted EPO with combinations of IH4 or 10F7 and a 35 or 5 amino acid

linker, and were incubated with RBCs. Both forms of Targeted EPO containing 10F7 showed a high number of RBCs bound to the coated petri dish, significantly more than RBC adhesion caused by those containing IH4. The IH4 fusion protein with a 5 amino acid linker resulted in a significantly lower number of bound RBCs compared to the one with a 35 amino acid linker (Figure 2.5C,D). These results suggest that a long linker and a membrane-distal epitope on GPA increase the potential for Targeted EPOs to bridge the two surfaces in these experiments.
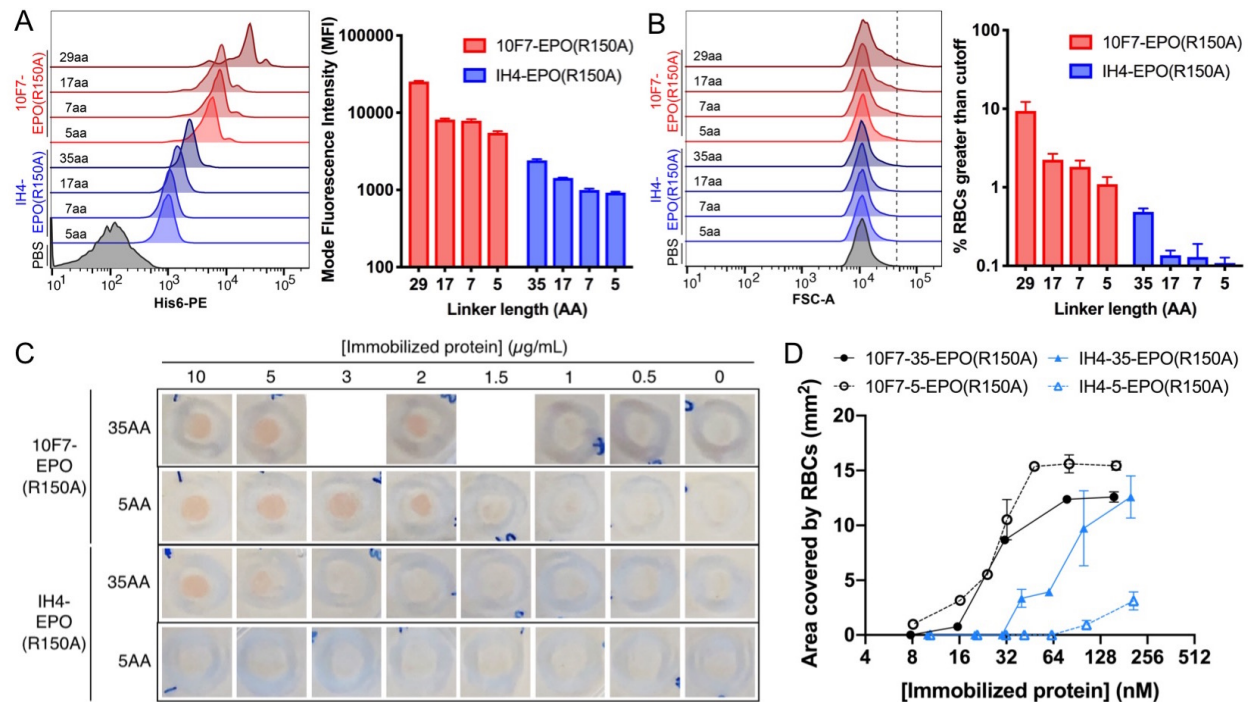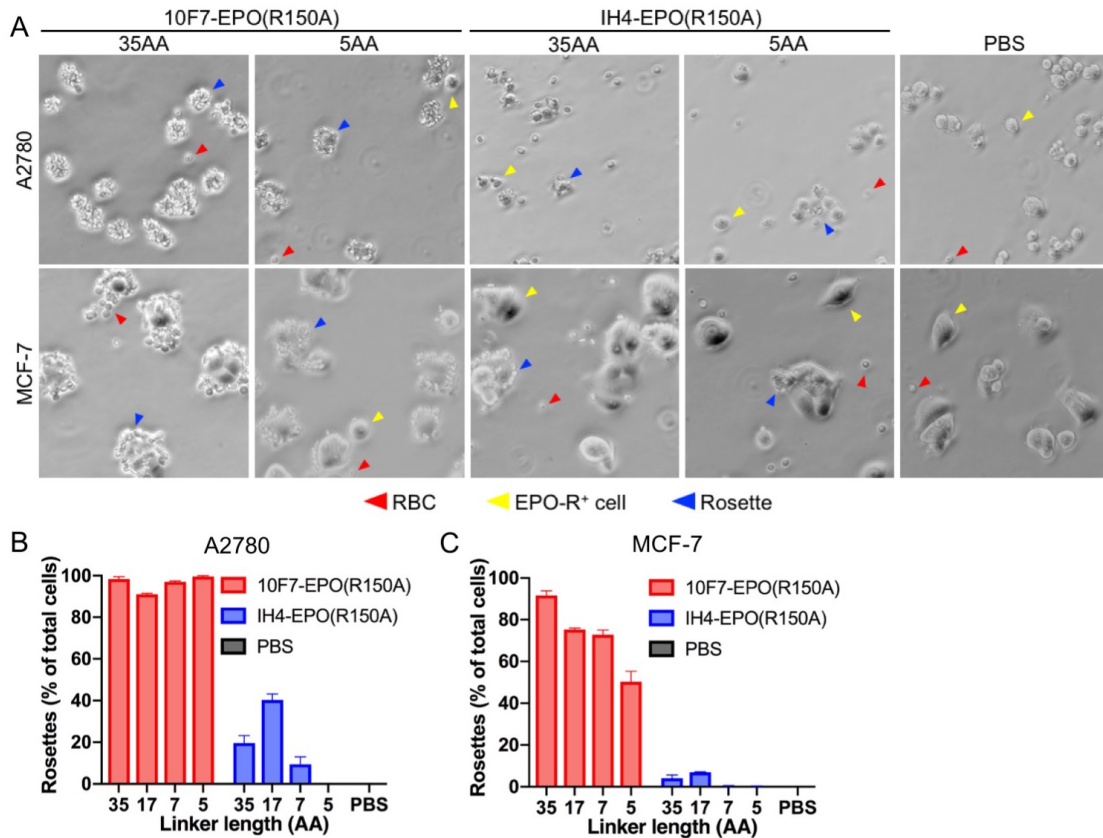

Figure 2.6: *Targeted EPO fusion proteins with short linkers and membrane-proximal GPA binding epitopes reduce potential for cell crosslinking. RBCs treated with Targeted EPO variants were incubated with A2780 or MCF-7, both of which express EPO-R. (A) Phase-contrast images (200× magnification) of RBC rosetting around EPO-R-bearing cells. (B,C) Frequencies of rosette formation, scored as tumor cells bound by three or more RBCs and shown as a percentage of total tumor cells. Data represent mean ± S.E.M.*

Consistently, forms of Targeted EPO with long linkers and a membrane-distal epitope on GPA mediate adhesion in vitro between RBCs and cells expressing EPO-Rs. I used a rosette assay for cell–cell interaction to address whether the anti-GPA/EPO fusion proteins promote adherence of RBCs to cells from the tumor lines A2780 and MCF-7, which express EPO-R [46, 47]. All forms of Targeted EPO containing the 10F7 scFv showed a high level of rosette formation,

with most, if not all, of the tumor cells binding to three or more RBCs. In contrast, rosette formation with Targeted EPOs containing the IH4 nanobody was much lower. Targeted EPO forms with shorter linkers generally induced less rosette formation. These effects were synergistic: the IH4-EPO(R150A) with a 5 amino acid linker showed essentially no rosette formation by either A2780 or MCF-7 cells (Figure 2.6 and Figure A.4).
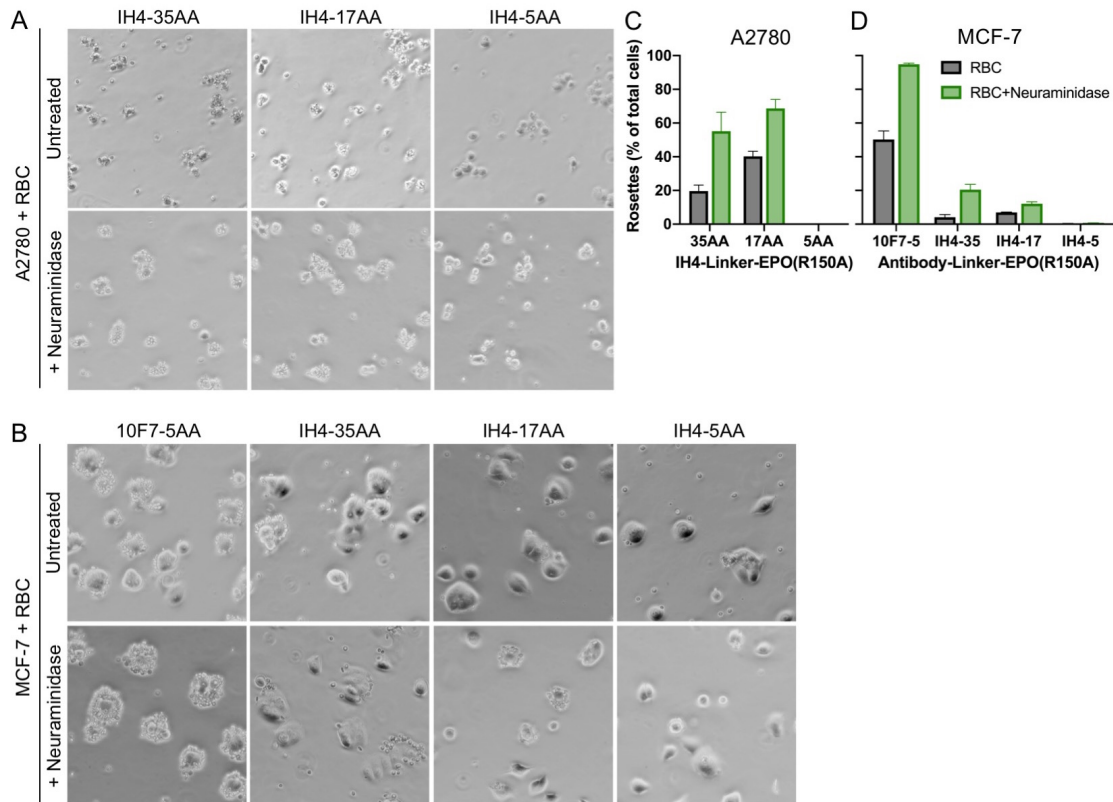
Figure 2.7: *N- and O-linked oligosaccharides of GPA reduce the ability of Targeted EPO fusion proteins to mediate cell crosslinking. Normal RBCs or RBCs treated with neuraminidase were treated with Targeted EPO variants, and were incubated with A2780 or MCF-7 cells, both of which express EPO-R. (A,B) Phase-contrast images of RBC rosetting around EPO-R-bearing cells. 200x magnification. (C,D) The frequency of rosette formation was scored as tumor cells bound by three or more RBCs, shown as a percentage of total tumor cells. Data represent mean $\pm$ S.E.M.*

Rosette formation mediated by Targeted EPOs is enhanced when the sialic acids of *N*- and *O*-linked oligosaccharides on GPA are removed. RBCs were treated with neuraminidase to cleave off sialic acids before incubating with a Targeted EPO and EPO-R-bearing tumor cells. These RBCs showed a higher level of rosette formation compared to their counterparts with intact oligosaccharides. The removal of sialic acids did not affect the relative levels of rosette formation among various forms of Targeted EPO. In both the normal and neuraminidase-treated RBCs,

IH4-EPO(R150A) with a 5 amino acid linker did not induce rosette formation (Figure 2.7).
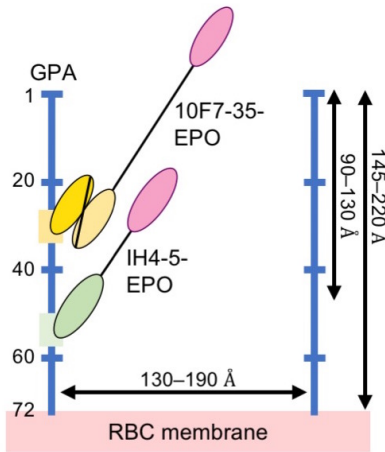
Figure 2.8: *Relative geometry and distances of GPA and the bound Targeted EPO on the RBC surface. The extracellular portion of GPA (blue line), showing the distance between the two adjacent GPA molecules, the height off the membrane, and 10F7-35-EPO and IH4-5-EPO bound at their respective binding epitopes ($_{34}YAATP_{38}$ and $_{52}YPPE_{55}$, respectively). The accessibility of GPA-bound IH4-5-EPO is sterically hindered by GPA.*

These results on antibody accessibility, RBC–petri dish binding, and rosette formation are consistent with a geometric model of the RBC surface in which GPA extends from the cell and shields the cell from interactions with other cells. RBCs have about 800,000 GPA monomers per 140 μm^2 surface area [34, 48], such that the average distance between GPA molecules is about 130–190 Å (Figure 2.8 and Figure A.5). The membrane-distal, N-terminal 45 amino acids of GPA include 17 sites for *N*- and *O*-linked glycosylation, with oligosaccharides that extend at least 20 Å from the peptide backbone (for the O-type blood group antigen) (Figure A.5) [49]. These glycans thus mask molecules that are closer to the membrane from outer molecules or other cells. It is thought that the N-terminal 45 amino acids of GPA are largely unstructured, and would be about 90–130 Å long in an extended conformation; the entire extracellular domain of GPA would be about 145–220 Å in an extended conformation [50]. By comparison, in an scFv–(SGGGS)–EPO fusion protein, the components would be 45, 15 and 51 Å end-to-end (a total of 111 Å) (Figure 2.8). Thus, contact between a GPA-bound IH4–(SGGGS)–EPO molecule and IgG antibodies, the surface of a petri dish, or EPO-Rs on another cell should be sterically very difficult.
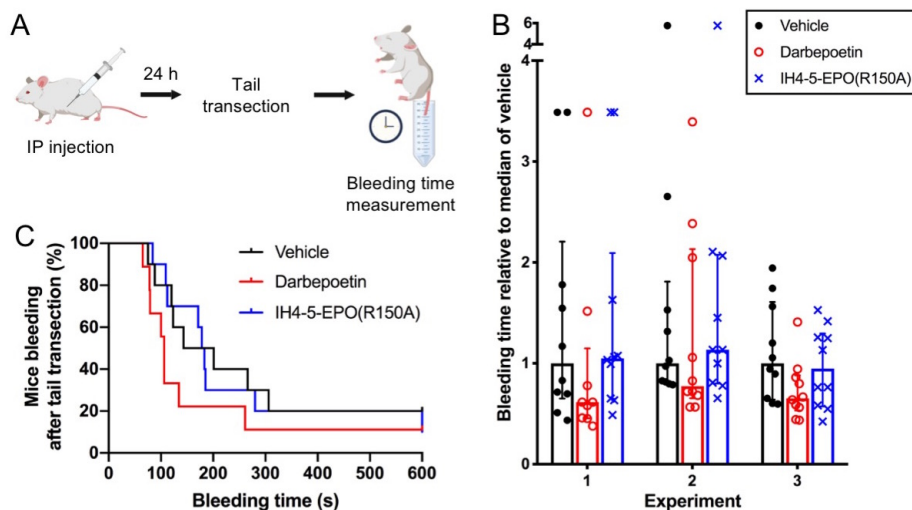
Figure 2.9: *Shortened tail bleeding time in mice treated with a non-targeted EPO but not those treated with IH4-5-EPO(R150A). (A) Experimental approach. Mice received a single i.p. injection of darbepoetin (50 pmol; 1.8 μg), IH4-5-EPO(R150A) (40 pmol; 2 μg), or vehicle. After 24 h, tail transection was performed, and bleeding times were recorded (see Methods). (B) Three experiments showing bleeding times relative to the vehicle control in huGPA-transgenic mice treated with vehicle, darbepoetin, or IH4-5-EPO(R150A). Data represent median with interquartile range. (C) Kaplan–Meier plot for Experiment 1 from the data shown in (B). Some images were created with BioRender.com.*

2.3.4 IH4-5-EPO(R150A) Does Not Shorten Tail Vein Bleeding Time in Mice

IH4-5-EPO(R150A) was tested for systemic effects on thrombosis in mice. HuGPA-transgenic mice were injected i.p. with vehicle, darbepoetin, 10F7-35-EPO(R150A) or IH4-5-EPO(R150A) (Figure 2.9A). EPO proteins were injected at a dose that resulted in similar effect in mice, as measured by reticulocyte counts. In mice treated with darbepoetin or 10F7-35-EPO(R150A), bleeding time was shortened by about 30%, whereas in mice treated with IH4-5-EPO(R150A) there was no change in bleeding time (Figure 2.9B,C and Figure A.1). These results indicate that IH4-5-EPO(R150A) does not enhance thrombotic side effects, in contrast to the non-targeted EPO form that acts on EPO-Rs on multiple cell types. When the data from these experiments were normalized to the vehicle controls of the same experiments and then combined, the difference between vehicle and darbepoetin (p = 0.0098) as well as IH4-5-EPO(R150A) and darbepoetin (p = 0.017) was significant. When data were compared across all experiments with dose groups including 10F7-35-EPO(R150A) and epoetin alfa, these molecules also showed a statistically significant decrease in bleeding time relative to the vehicle control (Figure A.1). These results are consistent with previous studies indicating that non-targeted EPO proteins reduce

bleeding time in rodents [51].

2.4 Conclusions

The design of a fusion protein must consider the spatial orientations of the target proteins to function while also minimizing unwanted interactions. Such considerations become important in AND-gate bifunctional protein systems that bind to different receptors abundant on multiple cell types. EPO-Rs are found on numerous cell types besides RBC precursors, including platelet precursors, vascular endothelial cells, liver cells, and white blood cells such as macrophages.

Here I addressed this problem for "Targeted EPO" molecules, fusion proteins that bind to GPA and weakly to EPO-R and are thus AND-gated to bind only to cells with both of these transmembrane proteins — namely late RBC precursors. The design goal is to activate EPO-R only on late RBC precursors as a result of initial binding to GPA, followed by binding in cis to EPO-R due to the high local concentration of EPO. I found that:

1. Some Targeted EPO proteins mediate adhesion between mature RBCs that express only GPA and other cells that express only EPO-R.

2. The position of the epitope on GPA and the length of the linker between the GPA binding element and the EPO element determines whether a Targeted EPO mediates cell adhesion.

3. The cell adhesion results can be interpreted with a structural model in which GPA extends from the RBC surface and limits how close another cell can approach. To mediate cell adhesion, the EPO element must extend beyond the membrane-distal end of GPA.

4. More broadly, cell–cell interaction may be regulated by abundant cell surface proteins that may sterically prevent close approach of other cells. This needs to be considered when constructing engineered cell interaction systems, such as CAR-T cells and bispecific antibodies.

A major challenge for synthetic biology is that artificial biological systems have the potential for undesired interactions that may lead to system failure. When we study natural systems, we

20

are generally not aware of all the potential behaviors they could have if poorly designed. My experiments illustrate geometric principles that natural systems may use in achieving specificity.

2.5   Materials and Methods

Cell Culture. FreeStyle 293-F, FreeStyle CHO-S, and CHO DG44 cells were obtained from Thermo Fisher Scientific (Waltham, MA), and were cultured in FreeStyle 293 Expression Medium, complete FreeStyle CHO Expression Medium, and complete CD DG44 (Thermo Fisher Scientific), respectively. Human erythroleukemia TF-1, human ovarian cancer A2780, and human breast cancer MCF-7 were obtained by ATCC (Manassas, VA). TF-1 was cultured in RPMI-1640 with 10% FBS, 100 U/mL penicillin, 100 U/mL streptomycin, and 2 ng/mL recombinant human granulocyte macrophage colony-stimulating factor (GM-CSF; PeproTech) unless specified otherwise. A2780 was cultured in RPMI-1640 with 10% FBS. MCF-7 was cultured in DMEM with 10% FBS and 0.01 mg/mL recombinant human insulin (PeproTech). 293-F, CHO-S, and CHO DG44 were cultured at 37°C in 8% CO_2 with shaking at 2.35 × g. TF-1, A2780, and MCF-7 were cultured at 37°C in 5% CO_2.

DNA Constructs for Protein Expression. The DNA sequences for 10F7 and EPO (wild-type and mutants) were from GenBank (accession no. KX026660–3). The DNA sequences for IH4, R18, and 1C3 were derived by reverse translating and codon optimizing (Integrated DNA Technologies) protein sequences adapted from US patent 9879090 [52], US patent 8900592 [53], and patent application WO1993024630 [54], respectively. The IH4 sequence was modified to include a point mutation (Phe80Tyr) and an additional amino acid (Thr118) in the framework regions 3 and 4, respectively, as the reported sequence of IH4 may have had fortuitous changes from the germline sequence or typographical errors. The IH4 antibody with the original sequence in the patent was denoted as IH4*. The DNA sequences for glycine–serine linkers of various lengths were codon-optimized for expression in mammalian cells.

Protein Expression and Purification. Transient expression was performed in 293-F and CHO-S cells using pSecTag2A or pOptiVEC plasmids, and stable expression was performed using CHO DG44 cells with pOptiVEC plasmids per the supplier's protocol. Protein expression was assayed by Western blotting cell supernatant using an anti-His$_6$-HRP antibody (Abcam). Stably

transfected cells were selected by hypoxanthine/thymidine (HT)-deficient CD OptiCHO medium (Thermo Fisher Scientific), and were subjected to one or more rounds of methotrexate (MTX; Sigma–Aldrich) genomic amplification as described before [33].

Proteins from transient transfection or stable pools were purified as follows. Supernatant was concentrated to 5–8 mL using a 10 kDa cut-off Macrosep Advance centrifugal device (Pall). Concentrated protein was bound to 0.5–1 mL of His60 nickel or HisTalon cobalt resin (Takara Bio) for 0.5 hr at 4°C while rotating in a 10-mL Pierce disposable column (Thermo Fisher Scientific), and was washed and eluted using His60 or HisTalon Buffer Set (Takara Bio) per the supplier's protocol. Cell supernatant and each purification fraction were analyzed by SDS–PAGE followed by Coomassie Blue staining. Eluted proteins were combined, desalted into endotoxin-free PBS (Teknova: 137 mM NaCl, 1.4 mM KH_2PO_4, 4.3 mM Na_2HPO_4, and 2.7 mM KCl, pH 7.4) using Econo-Pac 10DG columns (Bio-Rad), and concentrated to <1 mL using Macrosep Advance centrifugal devices. For in vivo experiments, contaminating proteins were further removed by size exclusion chromatography (SEC) on a Superdex 200 10/300 GL column (GE Healthcare) using the AKTA FPLC system (GE Healthcare) and endotoxin-free PBS as the running buffer. Sometimes, anion-exchange chromatography (AIEX) was performed on a HiPrep DEAE FF 16/10 column (GE Healthcare) using the AKTA FPLC system (GE Healthcare) prior to SEC. 20 mM Tris-HCl, pH 8.0 was used as the start buffer and salt concentration was elevated up to 1 M NaCl by a linear gradient. Desired protein fractions were combined and concentrated to <1 mL using Macrosep Advance centrifugal devices. Proteins were stored at 4°C throughout the described process, ultimately stored as aliquots at −80°C, and thawed once before use. Only endotoxin-free reagents were used.

Cell Proliferation Assays. TF-1 cells were seeded in a 96-well plate at 9.0×10^3 cells per well in 90 μL of RPMI-1640 with serum and antibiotics (no GM-CSF). The purified proteins were serially diluted by 10-fold (10^{-7} to 10^{-14} M) and added to the cells. After 72 hours, cell proliferation was measured by the CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega). Absorbance at 490 nm was read on a BioTek Synergy Neo HTS microplate reader. Reported data represent mean ± SEM of three replicates.

Measurement of Mouse Reticulocytes and Reticulated Platelets by Flow Cytometry. HuGPA transgenic FVB mice were generously donated by the Hendrickson Laboratory at Emory

University [55]. This strain underwent embryo re-derivation at Charles River Laboratories. The homozygous huGPA transgene is embryonic-lethal but heterozygotes are phenotypically normal, so a breeding colony was maintained with screening for huGPA at each generation. Transgene expression was measured as described before [33].

In experiments that measure efficacy of fusion and control proteins, five mice per dose group received a single intraperitoneal (i.p.) injection with saline, darbepoetin, or Targeted EPO in a 200 µL volume (diluted in saline or PBS) on Day 0. 1–5 µL of whole blood was collected by tail-nick in EDTA-coated tubes on days 0, 4, and 7 post-injection. Blood was analyzed immediately after collection by flow cytometry as described before [33]. Thiazole orange (Sigma–Aldrich) was used to stain residual RNA in reticulocytes and reticulated platelets, and anti-CD41-PE antibody (BD Pharmingen) was used to stain total platelets. A stock solution (1 mg/mL) of thiazole orange was prepared in 100% methanol and was diluted 1:5,000 in PBS to make a 2x working solution. Anti-CD41-PE antibody was diluted 1:500 in either the 2x working solution of thiazole orange for stained samples or PBS for gating thiazole orange-negative populations. 2 µL of whole blood was diluted 1:1,000 in 2 mL of PBS. Equal volumes (100 µL) of 2x working solution of anti-CD41-PE antibody, with or without thiazole orange and diluted whole blood, were mixed in a 96-well U-bottom plate and incubated for 30 min in the dark at 23°C. Fluorescence was measured on a LSRFortessa SORP flow cytometer equipped with an optional HTS sampler (BD Biosciences) using the following filter configuration: PE excitation, 561/50 mW; emission filter, BP 582/15; YFP excitation, 488/100 mW; emission filter, BP 540/25.

Preparation of Human Red Blood Cells. For FRAP, blood was obtained from healthy adult volunteers in accordance with the guidelines of, and approved by the Institutional Review Board of Beth Israel Deaconess Medical Center. Human blood was collected into HBSS++ containing 0.05% IgG-free BSA by venipuncture. RBCs were centrifuged at 5,000 x g at 23°C for 1 min, and were washed once with the same buffer. For flow cytometry and rosette assays, human whole blood collected in an ACD tube by venipuncture was purchased from Zen-Bio (Research Triangle Park, NC). RBCs were isolated by centrifugation at 5,000 x g at 23°C for 1 min, and washed twice with PBS-glucose (PBS, 5 mM glucose, pH 7.4).

Fluorescence Recovery After Photobleaching (FRAP). FRAP is a technique to measure membrane fluidity in mammalian cells such as RBCs, in which cells are incubated with a lipid-

soluble fluorescent dye to evenly label membranes and then a region of the membrane is bleached with a laser. The non-bleached dye from adjacent membrane will diffuse back into the bleached area at a certain rate that is a measure of the membrane fluidity. More specifically, RBCs (10% v/v in HBSS++ containing 0.1% IgG-free BSA) were incubated with lipid dye DiO (15 µg/mL) for 30 min at 23°C in the dark and washed once with HBSS++ containing 0.1% IgG-free BSA. They were then treated with GPA antibodies at 37°C for 15 min. The concentration of antibody was determined by its K_D to achieve 80% saturation of GPA. RBCs were imaged using a 60x objective on an Olympus BX62 fluorescence microscope, bleached with a Vector Controller Laser, and analyzed as described before [44, 56]. Reported data represent mean $\pm$ 95% confidence interval of 8–9 cells (experiment 1) or 16–20 cells (experiment 2). Data from each experiment were normalized to the buffer control in the same experiment to account for day-to-day variability, and the normalized values were combined across experiments. Mann–Whitney tests were performed to determine p-values for individual and combined data sets.

EPO Accessibility and RBC Crosslinking Measurement by Flow Cytometry. RBCs (Zen-Bio, Research Triangle Park, NC; 0.2% v/v in PBS-glucose) were incubated with fusion proteins for 1 hr at 4°C, washed once with PBS-glucose (PBS, 5 mM glucose, pH 7.4), incubated with anti-His$_6$-PE antibody (Abcam) for 30 min at 4°C, and washed once with PBS-glucose. Fusion protein concentration used was calculated based on the K_D of the monovalent 10F7 and IH4 antibody elements (95 nM and 33 nM, respectively) to achieve 80% saturation of GPA. The cells were analyzed by flow cytometry.

For the experiments in Figure 2.5A,B, the concentration of 10F7 fusion proteins was 385.4 nM in the binding reaction mixtures, while that of IH4 fusion proteins was 137.4 nM; the concentration of anti-His$_6$-PE antibody was 13.3 nM, and the concentration of available GPA epitopes on the RBCs was 8.8 nM (assuming 400,000 GPA epitopes per RBC) [57]. Cells were first incubated with the fusion protein for 1 hr at 4°C and washed, anti-His$_6$-PE antibody was added, cells were then incubated for 30 min at 4°C and washed again, and then they were analyzed by flow cytometry. Under these conditions, the fusion protein–GPA interaction is expected to reach equilibrium in a few minutes (Table A.2), while the anti-His$_6$-PE antibody–fusion protein interaction would likely reach equilibrium in tens of minutes for proteins free in solution; thus, a sterically-based reduction in on-rate of the anti-His$_6$-PE antibody could be observed.

24

Fluorescence was measured on a LSRFortessa SORP flow cytometer equipped with an HTS sampler (BD Biosciences) using the following filter configuration: PE excitation, 561/50 mW; emission filter, BP 582/15. Data were analyzed using FlowJo Data Analysis software (TreeStar, Ashland, OR). Reported data represent mean ± SEM of three replicates.

RBC–Petri Dish Binding Assay. On a petri dish, circles with a 0.5-cm diameter were drawn using a hydrophobic pap pen. 20 µL of fusion protein diluted in coating buffer (100 mM sodium carbonate/bicarbonate buffer, pH 9.6) was added to each circle and let sit for 1.5 hr at 23°C. Petri dish was washed three times with 10 mL of PBS, and incubated with 10 mL of blocking buffer (PBS, 0.02% Tween 20, 0.5% BSA) for 1 hr at 37°C. Petri dish was washed twice with 10 mL of PBS and once with PBS-glucose (PBS, 5 mM glucose, pH 7.4). 5.0×10^8 RBCs were added to the petri dish in 10 mL of PBS-glucose and incubated for 1 hr at 37°C. Petri dish was washed gently with 10 mL of PBS until background became clear. The area occupied by RBCs was estimated by measuring the number of pixels in Adobe Photoshop.

RBC Rosette Assay. The rosette assay is a particularly sensitive way to demonstrate the presence of surface receptors expressed at low levels [58]. We adapted this assay to study cell–cell adhesion mediated by fusion proteins that bind to RBCs. A2780 and MCF-7 cells were plated at 2.0×10^5 cells/mL in a 2 mL volume on a 12 x 12 mm coverslip (Electron Microscopy Sciences) in a well of a 6-well plate (Corning) and incubated overnight at 37°C in 5% CO_2. The following day, a mixture of RBCs (0.2% v/v) and fusion proteins was prepared in binding medium (RPMI-1640 without bicarbonate) and then incubated for 30 min at 23°C. In some experiments, RBCs were treated with α2-3,6,8,9-Neuraminidase A (New England BioLabs, Ipswich, MA) per the supplier's protocol before incubating with fusion proteins. Coverslips coated with cells were washed twice with PBS and transferred to a new 6-well plate. The RBC–fusion protein mixture (200 µL) was added on top of each coverslip and incubated for 1 hr at 37°C. Coverslips were gently washed twice with PBS. Rosettes were imaged on a Nikon Eclipse TE300 inverted phase-contrast microscope (Nikon, Melville, NY), using a Retiga EXi CCD camera (QImaging, Surrey, Canada) controlled by iVision 4.7 software (BioVision). Rosettes were defined as adherent cells bound by 3 or more RBCs and were shown as a percentage of at least 160 cells from at least 8 fields. Reported data represent mean ± SEM of at least two replicates.

Dose Determination for Mouse Tail Vein Bleeding Time Measurement. Dose levels

25

were chosen based on doses that were the minimum required to achieve an increase of reticulocytes of 10% of total RBCs relative to baseline 4 days after injection. For example, in a typical FVB mouse the fraction of reticulocytes is about 5%. In a mouse with a 5% reticulocyte baseline at day 0, an increase to 15% would be recorded as 10% relative to baseline. Four days is the point at which an increase in reticulocytes represents the maximal response. The potency of a given EPO-based molecule is a complex function of affinity for receptor, rate of receptor internalization and degradation during signaling, pharmacokinetics and distribution, and possible inflammatory effects that might antagonize erythropoiesis. Thus, different molecules were compared at functionally equivalent therapeutic doses. We found that doses of 1.8 µg of darbepoetin, 9 µg of 10F7-35-EPO(R150A), and 2 µg of IH4-5-EPO(R150A) had equivalent erythropoietic activity.

Mouse Tail Vein Bleeding Time Measurement. All mouse experiments were performed under the protocol IS00000723, approved by the Harvard Medical School IACUC. Experiments were performed in a blinded manner. Ten mice per dose group received a single i.p. injection with saline or test protein. On the next day, mice were anaesthetized. The tail was transected 3 mm from the tip, and the body of the mouse was placed on a heated pad over a 37°C water bath containing tubes with 50 mL of saline, such that 0.5–2 mm of the transected tail was placed in the saline within 2 seconds of transection. The bleeding time was recorded when bleeding stopped for one minute. Bleeding time longer than 10 minutes were recorded as 10 minutes.

A detailed description of the experiment is as follows. On day 0, mice were injected with a test protein (or vehicle). Because we were testing an unknown protein, it is important to also include mice that are injected with saline or PBS vehicle as a negative control, and EPO or darbepoetin as a positive control. Typically, 10 mice per dose group were used. The experiments were performed in a blind manner: one experimenter performed the injections of proteins into the mice, maintained the key, performed the injection of anesthetic, and then handed the mice in a random order to a second, blinded experimenter. The second experimenter performed the tail transections and measured the bleeding time. For most experiments, the tail transection and bleeding time was measured on day 1. In some experiments, the experiment was performed on day 4 so that changes in reticulocytes and reticulated platelets could also be measured. However, the advantage of performing the measurement 1 day after treatment is that after only 24 hours, the level of circulating red blood cells will not have changed, so effects on blood clotting are due

to direct effects on some element of the clotting system, and not due to changes in blood viscosity.

Tail transection was performed as follows. The mice were first anesthetized using anesthetics ketamine and xylazine. These anesthetics are chosen because they are thought to not affect blood clotting. Acepromazine was not used because it has the effect of reducing clotting. Mice were weighed, and mice were then injected with 120–160 mg ketamine /kg and 10–16 mg xylazine /kg of body weight. For older and heavier mice, sometimes an additional injection of about 25% of the first injection was required. After a mouse became unresponsive to a stimulus such as significant pressure to a hind foot, the mouse was placed on a heated pad on a platform over a water bath. The water bath was maintained at 37°C. 50-mL blue-cap tubes (Sarstedt) were filled with 50 mL of a solution of 0.85 to 0.9% NaCl that has been equilibrated to 37°C in a separate water bath. The animal was placed on a Chux pad for the transection.

A position on the tail that is 3 mm from the tip, not counting hair, was marked with a felt-tip pen using calipers. The tail was also inspected for signs of bruising that may be due to fighting, and data from such mice were discarded if the transected tail did not bleed at all. The decision to discard the data was always made in a blinded manner at the time of measurement. The tail was transected with a flat razor blade using a section of the blade that had not been used previously. Within 1–2 seconds, the transected tail was placed in a tube with pre-warmed saline, and then the body of the mouse was placed on the heated pad above the water bath. At the moment that the tail was transected by the blinded experimenter, the non-blinded experimenter started a timer. The body of the mouse was then positioned on the heated pad so that only the tip of the tail — about 0.5 to 2 mm — was in the saline, with the rest is in the air. When observing the bleeding tail, the tube was rotated so that the white stripe was behind the tail, providing contrast, and the room was well-lit. The rack holding the 50-mL tube was white or yellow to provide contrast. The bleeding time was recorded by noting cessation of bleeding, and then observing the submerged tail for up to one minute. If bleeding re-started within this minute, the first recorded time was not counted. Bleeding may stop and re-start several times. If the tail is still bleeding when 10 minutes have elapsed, the time was recorded as 10 minutes.

The median and mean bleeding times were calculated for each treatment group. Calculating the median has the advantage that extreme events, such as 10-minute time points, do not disproportionately contribute to the calculation. Data from each experiment were normalized to

27

the vehicle control in the same experiment to account for day-to-day variability. The normalized values were combined across experiments and Mann–Whitney tests were performed to determine the p-values of the combined data sets. Specifically, the mean of vehicle-treated controls was calculated for each experiment, and then individual data points for vehicle- and protein-treated animals were normalized to these mean values. Thus, the variability in vehicle control samples was preserved in the calculation of possible statistical significance. In the summary of these data shown in Figure A.1C, the normalized data from multiple bleeding time experiments were combined to estimate the overall magnitude and statistical significance of the treatment effects.

Chapter 3

Therapeutic Applications and Structural Analyses of Tissue-Protective Erythropoietin

3.1 Abstract

Erythropoietin (EPO) not only stimulates red blood cell (RBC) proliferation and maturation but also promotes blood coagulation, angiogenesis, and general protection from cell damage and death. The tissue-protective feature of EPO is thought to be mediated by an alternative receptor complex composed of an EPO receptor (EPO-R) and CD131, and has the potential to provide therapeutic benefits to a broad spectrum of diseases, in which the general property of tissue protection can be helpful. In this chapter, I studied the structure–function relationships of EPO in both hematopoietic and tissue-protective contexts to selectively exploit a subset of functions that are desired for a particular medical condition. I constructed EPO variants that exert one or both of RBC-proliferative and tissue-protective effects but not blood coagulation. A handful of EPO mutants were screened for their erythropoietic and tissue-protective activity in vitro in unfused form as well as in Targeted EPO form, characterized by an EPO mutant fused to an anti-GPA antibody fragment via a 5 amino acid linker (IH4-5-EPO(Mut)). Forms of Targeted EPO containing EPO mutations at the strong receptor-binding site, such as K45D and R150A, stimulated erythroid

cell proliferation but did not protect neuronal cells from cell death. Hence, these molecules are suitable for the treatment for anemia. Forms of Targeted EPO containing EPO mutations at the weak receptor-binding site, such as R103K and L108A, induced both erythroid cell proliferation and tissue protection, and therefore, would be useful for the treatment of high-altitude-related illnesses, high-altitude acclimatization, various military applications, and organ damage due to hypoxia from COVID-19, especially in patients on ventilators. Surprisingly, forms of Targeted EPO with the mutations R103K or L108A have an EC_{50} in the low femtomolar range, about 100,000-fold lower than other forms of Targeted EPO. I discuss possible mechanisms for this profound enhancement. I also built structural alignment models of the EPO–EPO-R–CD131 heterocomplex based on the structural homology of EPO to granulocyte-macrophage colony-stimulating factor (GM-CSF). These models, combined with the cell-based assay data, suggest potential binding modes of the EPO–EPO-R–CD131 heterocomplex as well as EPO residues that might be important for tissue protection. This work demonstrates that rational mutagenesis of a pleiotropic protein enables both the selective extraction of desired functions and silencing of undesired and potentially adverse effects.

3.2   Introduction

3.2.1   Pleiotropic Effects of Erythropoietin

Erythropoietin (EPO) is regarded as a "red blood cell (RBC)-producing" hormone, for its ability to stimulate RBC production in response to hypoxia. It inhibits apoptosis of late-stage erythroid precursor cells, such as BFU-E and CFU-E, and promotes their proliferation and maturation into the fully committed erythroid lineage. The kidneys of a healthy human adult constitutively produce EPO at a very low level, maintaining about 1–5 pM of circulating EPO under normoxic conditions to keep hemoglobin levels constant [59]. Under hypoxic stress or after massive blood loss, the body responds by increasing the production of EPO, which in turn, heightens the number of erythrocytes in circulation and allows for more efficient tissue oxygenation [60]. EPO stimulates RBC production by activating homodimeric EPO receptors (EPO-Rs), which are abundantly expressed in late erythroid precursor cells but then gradually disappear as the cells mature and get released into the bloodstream as reticulocytes and mature

RBCs [59].

Accumulating information suggests that EPO, like other cytokines and hormones, may be pleiotropic and perform several other biological functions in addition to hematopoiesis. Expression of functional EPO-Rs has been confirmed in many tissues other than erythroid precursors, such as endothelial cells, cardiomyocytes, and cells of the central nervous system, including the brain [61, 62, 63]. Deletion of EPO-Rs in mouse embryos resulted not only in impaired erythropoiesis, but also in developmental defects in the heart, the vasculature, and the brain [62]. The existence of functional EPO-Rs in non-hematopoietic tissues implies that EPO might activate EPO-Rs in different contexts to induce biological activities that are independent of erythropoiesis.

Non-hematopoietic functions of EPO that have been elucidated so far include aiding blood clotting, vascular growth, and tissue protection. These functions reshape our conception of EPO, suggesting that it may have been cleverly designed by nature as a "wound-healing" hormone, rather than a mere "RBC-producing" hormone. When an animal is wounded, the immediate response by the body should be to stop bleeding, to make more RBCs, to increase tissue oxygenation, and to ensure the survival of tissues until oxygen is supplied. EPO has displayed such non-hematopoietic effects in several experimental and clinical studies. For instance, recombinant human EPO, used for the treatment of anemia in patients with chronic kidney diseases and those undergoing chemotherapy, induced side effects that were pro-thrombotic, hypertensive, tumor angiogenic, and tumor proliferative [6, 7, 8, 9]. Several studies have shown that EPO promotes the expression and secretion of factors involved in blood clotting and hypertension by activating endothelial EPO-Rs, and that it enhances tumor growth by providing anti-apoptotic and proliferative signals via EPO-Rs on tumor cells and endothelial cells in the nearby vasculature [5]. EPO has also shown tissue-protective effects in several studies. It reduced neuronal cell damage and improved motor function in animal models of traumatic brain injury (TBI) or stroke [64, 65, 66, 67]. In human clinical trials, intravenous injections of high doses of EPO significantly reduced infarct size and serum markers of brain damage in acute ischemic stroke patients [68]. EPO treatment also resulted in a lowered mortality rate and improved neurological recovery amongst TBI patients [69]. Based on these observations, Burrill *et al.*, 2016 proposed that EPO should be considered a hormone that mediates the body's response to hemorrhage, rather than hypoxia [33]. This view underlies my interpretation of data in this chapter.

Due to its hematopoietic and tissue-protective functions, EPO holds great promise as a novel therapeutic for various applications. EPO's hematopoietic activity has already been utilized for the past three decades to correct anemia in renal failure and cancer patients. Many experimental studies and clinical trials have been carried out to harness the tissue-protective effects of EPO for several different neurological indications. However, the pleiotropic nature of EPO also complicates its therapeutic translation. Depending on the diseases being targeted, only the desired functions should be exploited independent of other unnecessary effects to ensure drug safety; while some functions may be beneficial, others may be useless or even deleterious in the treatment of a particular disease. Such separation of functions can be implemented by targeting EPO to the specific cell types that are involved in activating desired effects, by silencing undesired functions, or by choosing a dose at which only the desired functions are effective. While the hematopoietic action of EPO has been well studied, mechanistic and structural understanding of tissue-protective action of EPO is still lacking, and therefore, limits our capability to custom-design drugs that perform only the desired functions.

3.2.2   Role of EPO-R–CD131 Heterocomplex in Tissue Protection by EPO

Pleiotropic effects of a cytokine can arise by several mechanisms, such as acting on different cell types or cells at different stages of development, inducing signaling for differing durations of time, and binding to different receptor complexes [70]. Several studies suggest that EPO's tissue-protective effects may be mediated by a receptor complex that is different from the canonical EPO-R homodimer [64, 71]. In erythroid precursors, where EPO-Rs are most abundantly expressed (about 1000 copies per cell in CFU-E) [12], EPO binds to homodimeric EPO-Rs and induces the dimerization and autophosphorylation of JAK2, which then phosphorylates tyrosine residues at the cytoplasmic tail of EPO-Rs and activates downstream signaling pathways, most notably the JAK2–STAT5 pathway. As a result, anti-apoptotic genes are turned on and prevent erythrocyte precursors from dying so that these cells are able to proliferate and fully commit to the erythroid lineage [60, 62]. On the other hand, cytokine receptor common beta subunit (βcR), also known as CD131, in complex with EPO-R, has been implicated in tissue-protective EPO signaling [64, 71, 72]. CD131 is a signaling subunit for the receptors of granulocyte-macrophage colony-stimulating

factor (GM-CSF), interleukin-3 (IL-3), and IL-5. Although CD131 alone cannot bind these cytokines, it is essential for the high affinity binding of ligands to the receptor complexes composed of CD131 and cytokine-specific alpha subunits, as well as for ligand-induced signal transduction [73, 74]. Expression of both EPO-R and CD131 has been verified in various tissues in the central and peripheral nervous systems, retina, heart, kidney, muscle, and endothelium [62].

Engagement of CD131 in EPO signaling has been shown at a functional level in vitro and in vivo. EPO induced phosphorylation of CD131 in the UT-7 cell line within 1–30 minutes after stimulation, suggesting that CD131 directly responds to EPO binding and triggers downstream signaling [72]. The addition of neutralizing antibodies for EPO-R or CD131 ablated the anti-apoptotic effect of a tissue-protective variant of EPO that lacks erythropoietic activity (carbamylated EPO; CEPO) in a neuronal cell line [75]. The same molecule, CEPO, restored motor function after spinal compression in wild-type mice but not in CD131 knockout mice [64]. These results suggest that EPO may signal via EPO-R–CD131 to induce tissue-protective effects in vitro and in vivo.

Although several studies support tissue protection by EPO, EPO-R, and CD131, many of them do not address how they interact to form a signaling complex for tissue protection at a biochemical, biophysical, or structural level. The only biophysical study done so far claims that EPO does not induce any interaction between EPO-R and CD131, rejecting previous hypotheses and experimental results in support of CD131 as an alternative receptor for EPO [76]. However, this study used the soluble, extracellular domains of EPO-R and CD131, and relative concentrations of interacting partners were not reasonable enough to come to a conclusion. Such a lack of information about the protein structures and binding kinetics of EPO, EPO-R and CD131 limits our understanding of receptor–ligand interaction, which is crucial for the rational design of protein therapeutics.

3.2.3 Tissue-Protective EPO in Research and Clinical Studies

Recombinant human EPO has previously shown to improve neurological outcomes in patients or animal models with brain damage caused by TBI, stroke, and multiple sclerosis [68, 69, 77, 78]. In rats with mild TBI, administration of EPO led to a significant reduction in lesion size [77]. In patients who had severe closed brain injuries from road traffic crashes, those who

received EPO treatment showed a lower mortality rate and better neurological function [69]. In multiple sclerosis patients, EPO treatment improved motor and cognitive function [78]. In stroke patients, EPO treatment reduced infarct size and serum S100β levels, and improved stroke scale scores [68]. In all of these studies, high doses of EPO were administered to patients or animals within 6 hours after the injury or onset of symptoms for 3–7 days. In some studies in which patients were given insufficient doses or a single injection of EPO, the treatment did not have significant protective effects on brain damage [79]. However, subsequent studies done on stroke patients using high doses of EPO reported that the treatment increased the incidences of adverse events, including deaths [80].

In an effort to develop safer EPO drugs for treating neurological conditions, several groups have engineered EPO derivatives that provide tissue protection without its other effects, including erythropoiesis (Table B.1). For example, carbamylated EPO ("CEPO"), in which all the lysine residues are modified with a carbamoyl group ($-CONH_2$) and converted to homocitrulline, completely loses erythropoietic activity in vitro and in vivo due to extensively altered receptor contacts. Despite such changes, this molecule retained tissue-protective effects in vitro (defined by preventing apoptosis of cultured neurons subjected to hypoxic stress), and was able to protect neuronal cells from stroke, spinal cord compression, and diabetic neuropathy in rat models [71]. The observations that such a broad chemical modification did not affect the tissue-protective effects of EPO suggest that a region of EPO that did not contain any lysine residues may include a potential recognition site for the tissue-protective EPO receptor complex [81]. Helix B and the AB loop of EPO do not have lysine residues, and small peptide fragments from these two regions maintained tissue-protective effects despite loss of erythropoietic effects [81, 82]. An 11-mer peptide fragment was derived from helix B by combining the amino acids that are at the aqueous face of helix B, mimicking its solvent-exposed surface structure [81]. This peptide, named ARA290, was also shown to be tissue-protective but not erythropoietic, and has undergone clinical trials for several neurological conditions [81, 83, 84, 85, 86]. In another study, site-directed mutagenesis of EPO revealed mutants that disrupt erythropoietic activity but not tissue-protective activity. Several mutants in helix A and C, such as S104I, exhibited significantly weakened or completely abolished erythropoietic activity, and showed comparable tissue-protective effects to wild-type EPO [87]. Taken together, these studies suggest that the receptor-interacting epitopes

might be different for erythropoietic and tissue-protective effects, and demonstrate that the two effects may be separated by distorting or removing certain parts of EPO.

3.2.4 Limitations in Therapeutic Translation of Tissue-Protective EPO

While several studies have shown that EPO and its derivatives may be effective therapeutics for various neurological indications and TBI, their poor pharmacokinetic properties and high dose requirements have likely led to clinical failures. EPO is about 30.4 kDa with a short serum half-life of about 8 hours after a single intravenous injection in humans [12]. Aforementioned EPO derivatives are similar in size to or smaller than wild-type EPO, and thus, have serum half-lives of 8 hours or shorter. The peptide ARA290, for instance, is about 1.3 kDa and has a serum half-life of about 20 minutes after a single subcutaneous injection and about 2 minutes after a single intravenous injection in healthy humans [88]. Such short half-lives necessitate frequent injections or high doses of drugs. Weaker receptor interactions and the need to cross the blood–brain barrier (BBB) also increase the dose requirement. EPO binds to homodimeric EPO-R with high affinity (K_D = 0.1–1 nM) [12, 16, 59] while it binds to the tissue-protective receptor complex with about 100-fold lower affinity [89]. This implies that higher doses of EPO may be needed for tissue-protection than for RBC production. Additionally, in the context of tissue protection of the central nervous system, such as the brain, the amount of EPO available for target tissues is restricted by its efficiency crossing the BBB. Although EPO and its derivatives are much larger than the cut-off (400–600 Da) for BBB permeability by passive lipid-mediated transport [90], several studies have shown that EPO is still able to pass through [68, 91, 92, 93]. Mice that received a single intravenous injection of EPO had significantly elevated EPO levels in the brain, and stroke patients who received an intravenous infusion of EPO had 60–100 times higher levels of EPO in the cerebrospinal fluid than untreated patients [68, 91]. While the mechanism of how EPO crosses the BBB is not clear, it has been shown in rodents that only a tiny fraction of systemically injected EPO gets into the brain by either active or passive pathway [92, 93]. One study showed that EPO crosses the BBB in the absence of any neural insult, potentially through EPO-R-mediated transcytosis on the capillary endothelial cells at the brain-periphery interface [92]. In another study, about 0.05–0.1% of intravenously injected EPO per gram of brain weight crossed the BBB

35

in mice at a similar rate to albumin, suggesting that EPO might cross the BBB via extracellular pathways mediated by diffusion through a leaky BBB [93]. Therefore, clinically relevant doses of EPO must be much higher to deliver to target tissues in the central nervous system. However, the amount of EPO that is allowed in clinical trials is limited by previous studies that were focused on anemia in renal or cancer patients. These studies showed that high doses of EPO increased the chances of thrombovascular incidences in chronic renal failure patients, and consequently, the maximum allowed dose of EPO in clinical settings has been restricted to a non-thrombotic "safe" level that was likely below the effective dose for tissue protection [6, 7, 8].

In this work, I rationally design EPO-based protein therapeutics for various applications in which the general property of tissue protection could be useful. The strength of different EPO activities was modulated by incorporating various EPO mutants into the chimeric activator design described in Chapter 2. This way, I aim to improve pharmacokinetics and therapeutic windows, to allow the administration of clinically relevant doses that are not limited by the known adverse effects of EPO drugs. Furthermore, these engineering strategies enable customized designs of various flavors of EPO therapeutics that protect neuronal cells from damage caused by various conditions, such as TBI, high-altitude-related illnesses, and neurodegenerative disorders.

3.3 Results and Discussion

3.3.1 Rationale for the Design of Tissue-Protective Forms of EPO

The goal of this chapter is to custom-design EPO therapeutics to accommodate different medical conditions by modulating EPO's erythropoietic and tissue-protective functions. The general strategy for designing such molecules was based on the concept of a "chimeric activator," introduced in Chapter 2. In short, a chimeric activator is a fusion protein composed of an antibody element that is targeted to specific cell types expressing antigens, a flexible peptide linker, and an activity element (cytokine or hormone) with a mutation that weakens its affinity for receptors. This way, the activity element acts only on target cells, mitigating unwanted effects via non-target cells [11, 33, 94]. The fine-tuning between the erythropoietic and tissue-protective effects of EPO was achieved by using different mutant proteins with different binding affinities for each of the responsible receptor complexes. By modulating the type and strength of mutations, I designed

three types of EPO fusion proteins, EPO-A, EPO-P, and EPO-AP (A: anemia; P: protection), each of which has a different subset of EPO functions (Figure 3.1 and Table 3.1). EPO-A only has an erythropoietic effect, and is designed as a replacement for current ESAs that carry lethal thrombotic side effects, as discussed in Chapter 2. It is intended for safer treatment of anemia in chronic kidney disease patients and chemotherapy patients. EPO-P only has a tissue-protective effect, and would be useful for treating neurodegenerative disorders as well as any conditions that accompany short-term hypoxic damage or other pro-apoptotic damage, such as TBI, strokes, and surgeries. EPO-AP has both erythropoietic and tissue-protective effects, and would be beneficial for patients suffering from high altitude-related illnesses as well as military personnel acclimating to high-altitude regions, seeking safe ways to enhance physical performance, and needing general protection from potential threats, such as chemical warfare. EPO-AP may also alleviate organ damages caused by hypoxia in COVID-19 patients (Table 3.1).

Table 3.1: *Three types of engineered EPO fusion proteins and their therapeutic applications.*

	EPO-A	EPO-P	EPO-AP
Molecules	IH4-5-EPO(Mut1)	IH4-5-EPO(Mut2)	IH4-5-EPO(Mut3)
Bioactivities	• RBC production	• Tissue protection	• RBC production • Tissue protection
Therapeutic applications	• Anemia due to chronic kidney failure or chemotherapy	• Traumatic brain injuries • Neurodegenerative diseases • Surgery enhancement	• High-altitude-related illnesses • Performance enhancement and tissue protection for military personnel • COVID-19-induced organ damage
Shared features	• Reduced pro-thrombotic side effects • Increased serum half-life • Reduced immunogenicity		

All of these molecules share common features that improve drug safety and pharmacokinetics. Fusing mutated EPO to the anti-glycophorin A (GPA) nanobody element, IH4, via a short 5 amino acid linker, as discussed in Chapter 2, not only increases the size of the molecule to above the renal filtration cut-off but also directs the fusion proteins to mature RBCs in circulation, extending serum half-life and reducing immunogenicity of the drug by inducing tolerance [28, 29, 33].
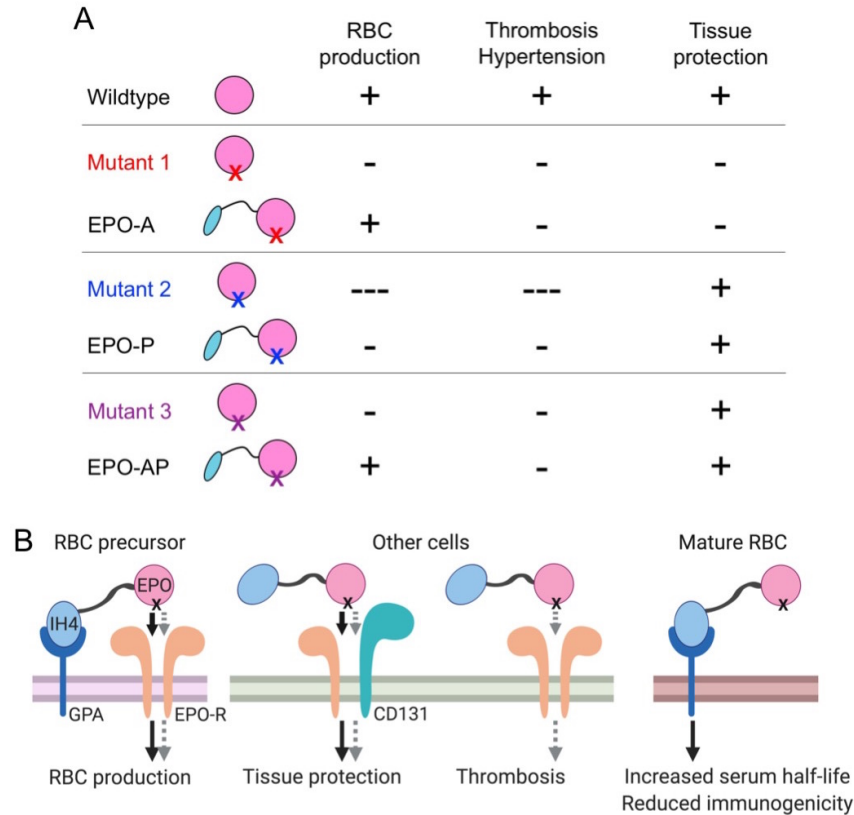
Figure 3.1: *EPO actions are mediated by different receptor complexes and can be modulated by implementing different mutations. (A) EPO mutants 1, 2, and 3 lack all or a subset of EPO activities. Their ability to promote RBC production can be rescued by fusing to an anti-GPA antibody fragment, depending on the strength of the mutation. (B) Surface receptor binding of IH4-5-EPO(Mut) fusion protein on different cell types. Fusion IH4 to EPO increases serum half-life and reduces immunogenicity of the therapeutic protein. Different mutations in EPO can be designed to weaken interaction with only one or both of the receptor complexes. Particular bioactivities of EPO can be selectively included or excluded from therapeutic designs. Some images were created with BioRender.com.*

EPO mutations in all three molecules generally weaken its affinity to homodimeric EPO-R to avoid undesired pro-thrombotic effects, triggered by homodimeric EPO-R signaling on non-target cells. The extent to which each of these mutations disrupts homodimeric EPO-R interaction is based on whether or not erythropoietic activity is desired. At the same time, it is important to ensure that the same mutation does not disrupt EPO binding to the EPO-R–CD131 complex when the tissue-protective activity of EPO is desired (Figure 3.1). Several EPO mutants were designed based on EPO mutagenesis studies previously done by other groups [87, 95]. Because the binding mode of EPO to EPO-R–CD131 is not elucidated, unlike that of EPO to homodimeric EPO-R, single point mutations were made in both of the known EPO-R contact regions (strong and weak interaction sides) (Figure 3.2). These EPO mutants were expressed and purified in an

unfused form (EPO(Mut)) and an antibody-fused form (IH4-5-EPO(Mut)), and were tested for erythropoietic and tissue-protective activities.

	Amino acid	Position	Mutate to
Strong Side	Lys	45	Asp
	Arg	150	Ala
Weak Side	Arg	14	Glu
			Gln
			Asn
	Tyr	15	Ile
	Arg	103	Lys
			Gln
			Ile
	Ser	104	Ile
	Leu	108	Ala
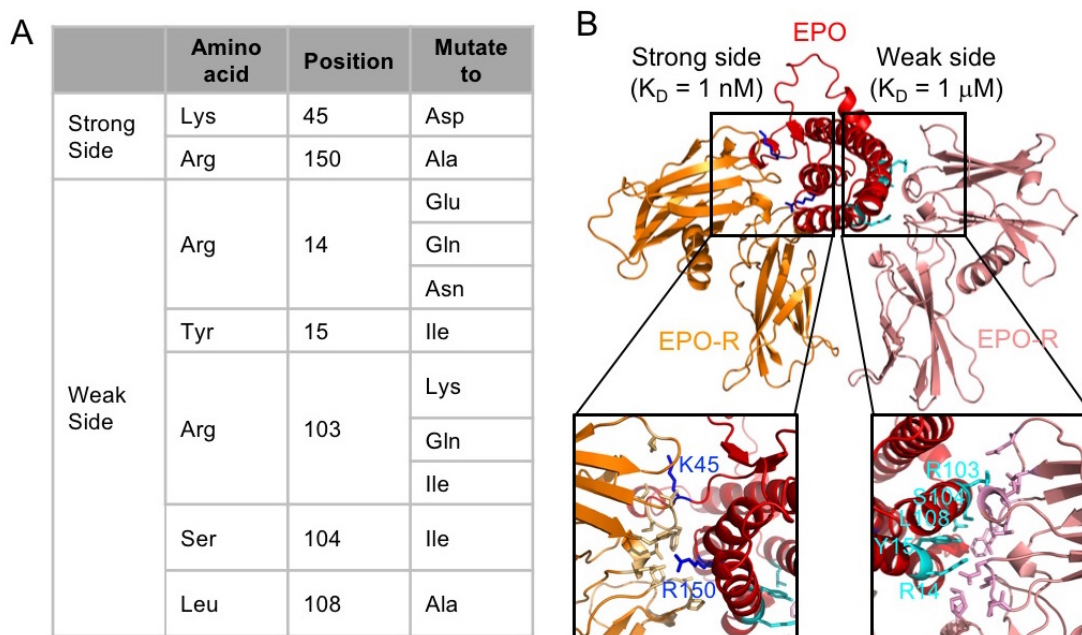
Figure 3.2: *(A) A list of EPO mutations studied in this chapter. (B) Structure of EPO in complex with EPO-R homodimer (PDB ID: 1EER) showing EPO residues that are mutated in this study. Note that these residues are located at both the strong and weak receptor-binding interfaces. Those at the strong and weak sides are shown in blue and cyan, respectively. EPO-R residues that are within 4 Å from these EPO residues are shown as sticks. EPO residues to be mutated make important contacts with EPO-R.*

3.3.2 Erythropoietic Activity of EPO Variants

The ability of different EPO mutants to promote RBC production was tested in vitro via TF-1 cell proliferation assays. The TF-1 cell line is of immature erythroid origin, expresses both EPO-R and GPA (1620 $\pm$ 140 and 3860 $\pm$ 780 molecules per cell, respectively) [11], and requires EPO, GM-CSF, or IL-3 for growth [96]. TF-1 cells that were starved of cytokines overnight were incubated with different EPO variants for 72 hr, and then their proliferation was measured by standard tetrazolium-based assay. Wild-type EPO (epoetin alfa, Epogen®, Amgen) and the hyperglycosylated form of EPO (darbepoetin alfa, Aranesp®, Amgen) exhibited EC$_{50}$ values of about 0.1 nM and 1 nM, respectively. EPO mutations on the strong side reduced the activity of unfused EPO by about 120- to 3400-fold relative to epoetin alfa. When these mutants were fused to the IH4 nanobody, their activities were rescued by about 180- to 340-fold relative to

unfused mutants, showing comparable activity to epoetin alfa and darbepoetin (Figure 3.3A). All unfused EPOs with mutations on the weak side did not show any activity at concentrations ranging from 10^{-14} to 10^{-7} M, except for EPO(R103K) (Figure 3.3A,B and Figure B.1). EPO(R103K) had a comparable EC_{50} value to epoetin alfa but about two-fold lower efficacy (E_{max}).

A

	Amino acid	Position	Mutate to	*In vitro* erythropoietic activity					
				EPO			IH4-5-EPO		
				N	Log(EC$_{50}$) (M) ± S.D.	EC$_{50}$ relative to epoetin alfa	N	Log(EC$_{50}$) (M) ± S.D.	EC$_{50}$ relative to epoetin alfa
Control	Epoetin alfa (Epogen®)			21	-10.22 ± 0.32	1	0	N.D.	N.D.
	Darbepoetin (Aranesp®)			10	-9.06 ± 0.27	14.20	0	N.D.	N.D.
Strong Side	Lys	45	Asp	4	-6.68 ± 0.65	3425.96	2	-9.21 ± 0.49	10.04
	Arg	150	Ala	8	-8.14 ± 0.06	119.32	9	-10.41 ± 0.29	0.64
Weak Side	Arg	14	Glu	1	N.A.	N.A.	2	N.A.	N.A.
			Gln	1	N.A.	N.A.	1	N.A.	N.A.
			Asn	1	N.A.	N.A.	1	N.A.	N.A.
	Tyr	15	Ile	2	N.A.	N.A.	1	N.A.	N.A.
	Arg	103	Lys	6	-9.93 ± 0.40	1.94	3 8	EC$_{50}$: -14.74 ± 1.74 IC$_{50}$: -8.92 ± 0.39	EC$_{50}$: 2.99x10^{-5} IC$_{50}$: 19.78
			Gln	2	N.A.	N.A.	2	N.A.	N.A.
			Ile	1	N.A.	N.A.	2	N.A.	N.A.
	Leu	108	Ala	5	N.A.	N.A.	2 5	EC$_{50}$: -14.64 ± 0.01 IC$_{50}$: -9.53 ± 0.22	EC$_{50}$: 3.85x10^{-5} IC$_{50}$: 4.90

B

Figure 3.3: *Erythropoietic activities of several EPO variants in vitro. (A) Table summarizing the mutations at strong and weak sides of EPO studied in this chapter and their in vitro erythropoietic activities measured by TF-1 cell proliferation assays. (B) Typical TF-1 cell proliferation assay was performed to compare the stimulation of cell proliferation by darbepoetin (hyperglycosylated form of EPO) and two EPO mutants, EPO(R103K) and EPO(L108A), in unfused and fused forms. EPO(R103K) and EPO(L108A) fused to the IH4 nanobody via a 5 amino acid linker rescued or enhanced the activity of EPO mutants alone. Their targeted erythropoietic activity shows bell-shaped dose response curves with very high potency. Data represent mean ± S.E.M. of three replicates. N.A. = No activity. N.D. = Not determined.*

The dose-response curve of weak-side mutant EPOs fused to IH4 has three unusual features. First, when EPO(R103K) was fused to IH4, the potency of the fusion protein was enhanced by five orders of magnitude relative to the unfused mutant and to other EPO fusion proteins. The EC_{50} is about 1–5 fM (Figure 3.3A,B). Among the weak-side mutants that lack erythropoietic activity, only

EPO(L108A) exhibited targeted erythropoietic activity upon fusion to IH4 and also has an EC_{50} of about 1–5 fM, while the others remained inactive even after fusion (Figure 3.3A,B and Figure B.1). Both of these mutations have the effect of reducing the size of a side chain and thus removing a contact with the receptor, but do not create steric hindrance for binding. Secondly, the maximal stimulation of TF-1 cell proliferation by IH4-5-EPO(R103K or L108A) is about 1/3 to 1/2 of the stimulation seen with wild-type EPO and other EPO fusion proteins (Figure 3.3B). Lastly, the dose-response curve of IH4-5-EPO(R103K or L108A) is bell-shaped, with stimulation falling off at about 0.1–0.5 nM, whereas fusion proteins containing strong-side mutants (K45D and R150A) showed standard sigmoidal dose-response curves (Figure 3.3B and Figure B.1). Hypotheses to explain each of these results are proposed as follows.

First, two possible mechanisms could explain the extremely high potency of IH4-5-EPO(R103K or L108A).

1. The attachment of the fusion protein to GPA could prevent receptor-mediated endocytosis and degradation of the signaling protein. GPA is attached to a stable actin cytoskeleton that may preclude internalization [97, 98].

2. The fusion protein may form a highly stable complex with GPA and one copy of EPO-R via the strongly interacting side of EPO, but the interaction with the second EPO-R to form a complete signaling complex may be very weak and dissociate rapidly due to a mutation. The interaction may last long enough to phosphorylate a subset of the tyrosine residues important in signal transduction into the nucleus but may not be long enough to phosphorylate those involved in signaling to the clathrin system for receptor-mediated endocytosis.

The former model is contradicted by the fact that I do not observe highly potent activity with other Targeted EPO fusion proteins. However, the underlying concept of this model may explain how such a small amount of fusion protein can trigger cell proliferation without being cleared by non-receptor-mediated endocytosis that routinely occurs as part of the membrane turnover. Additionally, the ability of IH4-5-EPO(R103K or L108A) to stimulate cell proliferation at very low concentrations suggests that only a small number of complete receptor–ligand complexes are required and sufficient to induce proliferative signaling. In this experiment, each well of a

41

96-well plate initially contained 9,000 cells and about 6.022×10^4 copies of fusion proteins (when 1 fM was given), thereby resulting in a molar ratio of 1:1.67 between the number of cells and that of fusion proteins per well. Assuming that cells and proteins are evenly distributed, only 1 to 2 copies of fusion proteins on average would bind to receptors on each cell at a given time and stimulate cell proliferation. This extraordinary feature of IH4-5-EPO(R103K or L108A) may be possible because complete dissociation of the fusion protein from GPA and EPO-R occurs at a very low rate. IH4-5-EPO(R103K or L108A) practically has a bivalent targeting element consisting of IH4 and the strong side of EPO. Having two target sites increases the net affinity ("avidity effect") and reduces the overall dissociation rate of the fusion protein from its targets. When one part of the fusion protein, IH4 for example, dissociates from its target site, GPA, the other part of the fusion protein, EPO in this case, may remain bound to its target site, EPO-R, and force IH4 to stay in proximity such that re-binding can occur rapidly.

Secondly, the failure to form a long-lasting stable signaling complex could explain the low maximal stimulation of cell proliferation. The fusion protein may form a tight complex with GPA and one copy of EPO-R but its binding to the second copy of EPO-R is very weak ($K_D > 1$ μM). Rapid dissociation of the second EPO-R may interfere with full activation of EPO-R, thereby explaining the activity profile resembling that of a typical partial agonist, with less than half of the maximal efficacy observed with wild-type EPO and other fusion proteins.

Lastly, the bell-shaped dose-response curve of IH4-5-EPO(R103K or L108A) can be explained by receptor saturation that has been observed in other systems that require more than two receptors for signaling [99, 100, 101]. The fusion protein may saturate monomeric EPO-R in a 1:1 stoichiometry and block the formation of a complete signaling complex consisting of homodimeric EPO-R. Calculations based on the dissociation constant (K_D) of EPO for EPO-R on the strong side and the concentrations of the applied fusion protein and EPO-R show that the addition of 0.1 nM of fusion protein would result in about 9.1% of EPO-R being bound by the fusion protein via the strong interaction side. Similarly, adding 1 nM and 10 nM would lead to about 50.0% and 90.8% of EPO-R being occupied by the fusion protein, respectively. This is consistent with the experimental data showing that the activity of IH4-5-EPO(R103K or L108A) begins to fall off at 0.1–0.5 nM and is completely lost at >5–10 nM (Figure 3.3B).

These results indicate that IH4-5-EPO fusion proteins composed of the K45D, R150A,
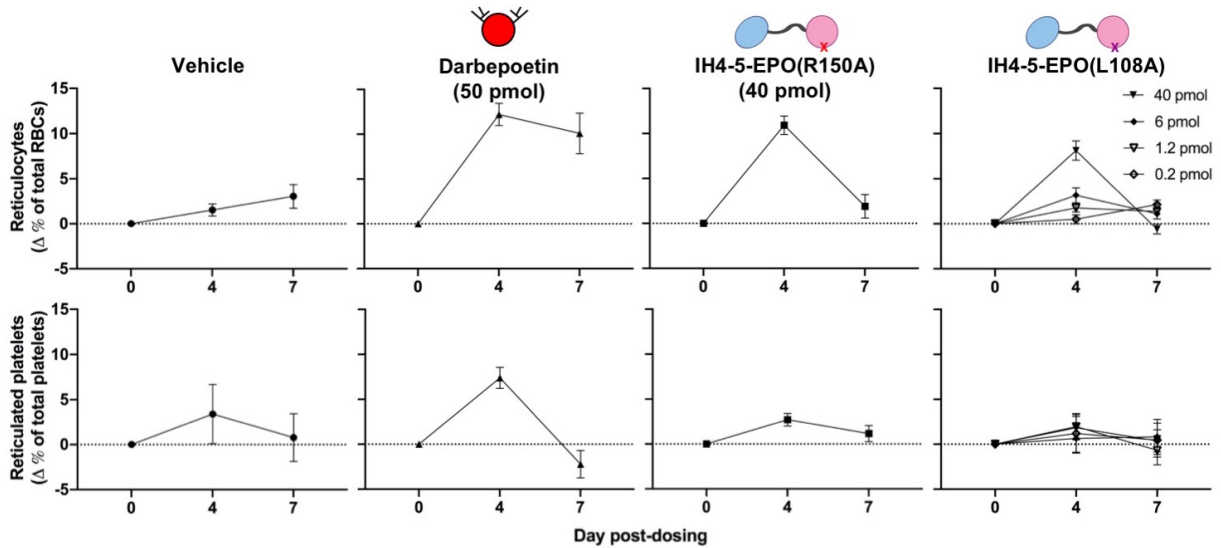
Figure 3.4: *Erythropoietic activities of IH4-5-EPO(L108A) in vivo. IH4-5-EPO(L108A) specifically stimulates RBC production and not platelet production in transgenic mice that express human GPA on their RBCs. Erythropoietic response to IH4-5-EPO(L108A) in mice is dose-dependent. Data represent mean ± S.E.M of five mice per dose group.*

R103K, or L108A mutants induce erythroid cell proliferation in vitro, and therefore, are promising candidates for EPO-A or EPO-AP. IH4-5-EPO(R150A) had been tested extensively for its targeted erythropoietic activity as well as reduced systemic pro-coagulative effects in mice, as discussed in Chapter 2. Here, one of the weak-side mutant fusion proteins, IH4-5-EPO(L108A), was further tested for its targeted erythropoietic activity in transgenic mice expressing human GPA. Mice received a single intraperitoneal (i.p.) injection of saline, darbepoetin (50 pmol = 1.8 µg), IH4-5-EPO(R150A) (40 pmol = 2 µg), or various doses of IH4-5-EPO(L108A) (40 pmol = 2 µg; 6 pmol = 0.3 µg; 1.2 pmol = 0.06 µg; 0.2 pmol = 0.01 µg). Target cell specificity and drug efficacy were measured by staining for reticulocytes and reticulated platelets in blood samples on Days 0, 4, and 7. Both the reticulocyte and reticulated platelet levels remained at baseline (Day 0) throughout the experiment in the saline-treated mice, but increased significantly in mice treated with darbepoetin, a control for the non-targeted form of EPO. Mice treated with IH4-5-EPO(R150A) had significantly elevated reticulocyte levels (10.9%) that were comparable to those in the darbepoetin-treated mice (12.2%) on Day 4, but did not have increased reticulated platelet counts. IH4-5-EPO(L108A) behaved similarly to IH4-5-EPO(R150A), in that it elevated reticulocyte levels but not reticulated platelet levels. IH4-5-EPO(L108A) induced reticulocyte responses in a dose-dependent manner. 40 pmol and 6 pmol induced 8.13% and 3.16% increases in reticulocyte counts on Day 4 relative

to Day 0, respectively, while lower doses (1.2 pmol and 0.2 pmol) did not have significant effects (Figure 3.4 and Figure B.2). It would be interesting to test a wider range of concentrations to see how the in vitro bell-shaped dose-response curves translate in animal studies.

3.3.3   Tissue-Protective Activity of EPO Variants

EPO mutants were further tested for their tissue-protective effects in cell-based assays. Those that showed targeted erythropoietic activity were evaluated for their potentials for either EPO-A or EPO-AP. Those that lacked erythropoietic activity were considered for EPO-P. The ability of a fusion protein to protect cells from a toxic damage was measured in vitro by estimating the number of surviving cells after treatment with EPO and a toxic agent. In more detail, SH-SY5Y, a neuroblastoma cell line that expresses both EPO-R and CD131, was pre-treated with EPO or an EPO variant for 24 hr, and then received 100 µM of cobalt chloride ($CoCl_2$), which mimics hypoxia by driving cellular responses initiated by hypoxia-inducible factor-alpha (HIF-1α). 24 hr later, the number of viable cells were measured by standard tetrazolium dye-based assays. The optimal cell density and concentration of $CoCl_2$ were determined experimentally to cause about 40% cell death (Figure B.3).

Results showed that wild-type EPO (EPO(WT)) and EPO(S104I) protect neuroblastoma cells from $CoCl_2$ insult. This is consistent with the previous results that EPO(WT) and EPO(S104I) protected primary neurons from NMDA-induced excitotoxicity (Figure 3.5A,C) [87]. I also found that darbepoetin, hyperglycosylated EPO, provides some level of tissue protection (Figure 3.5B). This indicates that the additional oligosaccharides did not completely disrupt the receptor–ligand interaction essential for tissue protection. Among the EPO mutants that stimulated TF-1 cell proliferation upon fusion to IH4, two weak-side mutants, EPO(R103K) and EPO(L108A), exhibited tissue-protective effects in both unfused and fused forms (Figure 3.5F–I). In contrast, two strong-side mutants, EPO(K45D) and EPO(R150A), did not protect cells from $CoCl_2$-induced cell death (Figure 3.5D,E). Four-parameter fits of these data did not give accurate EC_{50} values because the read-outs did not reach saturation. Despite this caveat, four-parameter fits provided rough estimates for the potency of each mutant. The EC_{50} values of EPO(WT) and EPO(S104I) were about 2 nM, and darbepoetin about 290 nM (ambiguous fit). The EC_{50} values of EPO(R103K) and

44
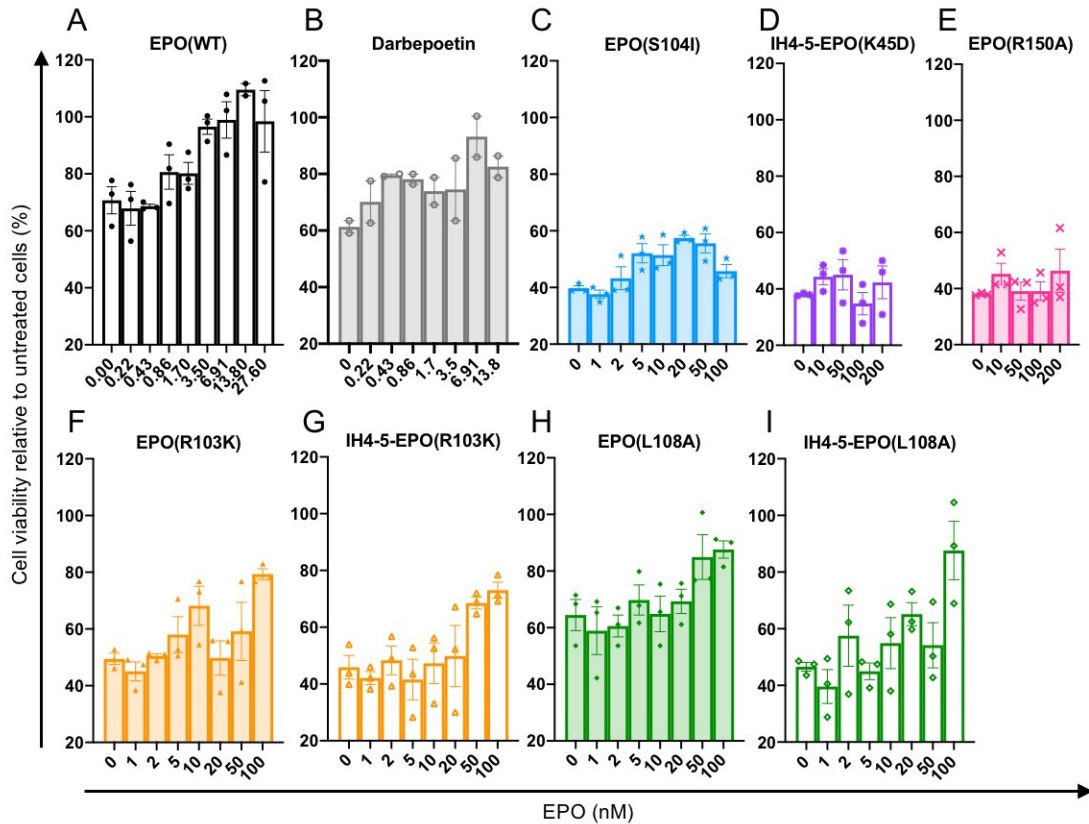
Figure 3.5: *Tissue-protective activities of several EPO variants in vitro. Viable cell counts were measured after treating cells with EPO variants and CoCl$_2$. (A,B) Wild-type EPO (EPO(WT)) and darbepoetin protect cells from cobalt-induced cell death. (C) EPO(S104I), a positive control adopted from Gan et al., 2012, also stimulated tissue protection in this assay. (D,E) Strong-side mutants, EPO(K45D) and EPO(R150A), did not protect cells from dying. (F–I) Weak-side mutants, EPO(R103K) and EPO(L108A), had tissue-protective effects in both unfused and fused forms. Data represent mean $\pm$ S.E.M. of two or three replicates.*

EPO(L108A) were predicted to be around 30 nM.

Since the K45D and R150A mutants do not have significant tissue-protective activities, they remain good candidates for EPO-A, as extensively discussed in Chapter 2. IH4-5-EPO(R103K) and IH4-5-EPO(L108A) have both erythropoietic and tissue-protective activities, and therefore, hold potential for EPO-AP. Due to the complete lack of homodimeric EPO-R activation by its unfused counterpart, IH4-5-EPO(L108A) was chosen as a better candidate for EPO-AP, and will be further investigated in an animal model of TBI to corroborate its tissue-protective effects in vivo. The EPO mutants that completely lost erythropoietic activity in both unfused and fused forms, such as EPO(R14E/Q/N), EPO(Y15I), and EPO(R103Q/I), need to be tested for their tissue-protective activities to determine final candidates for EPO-P.

3.3.4 Structural Modeling of EPO–EPO-R–CD131 Heterocomplex
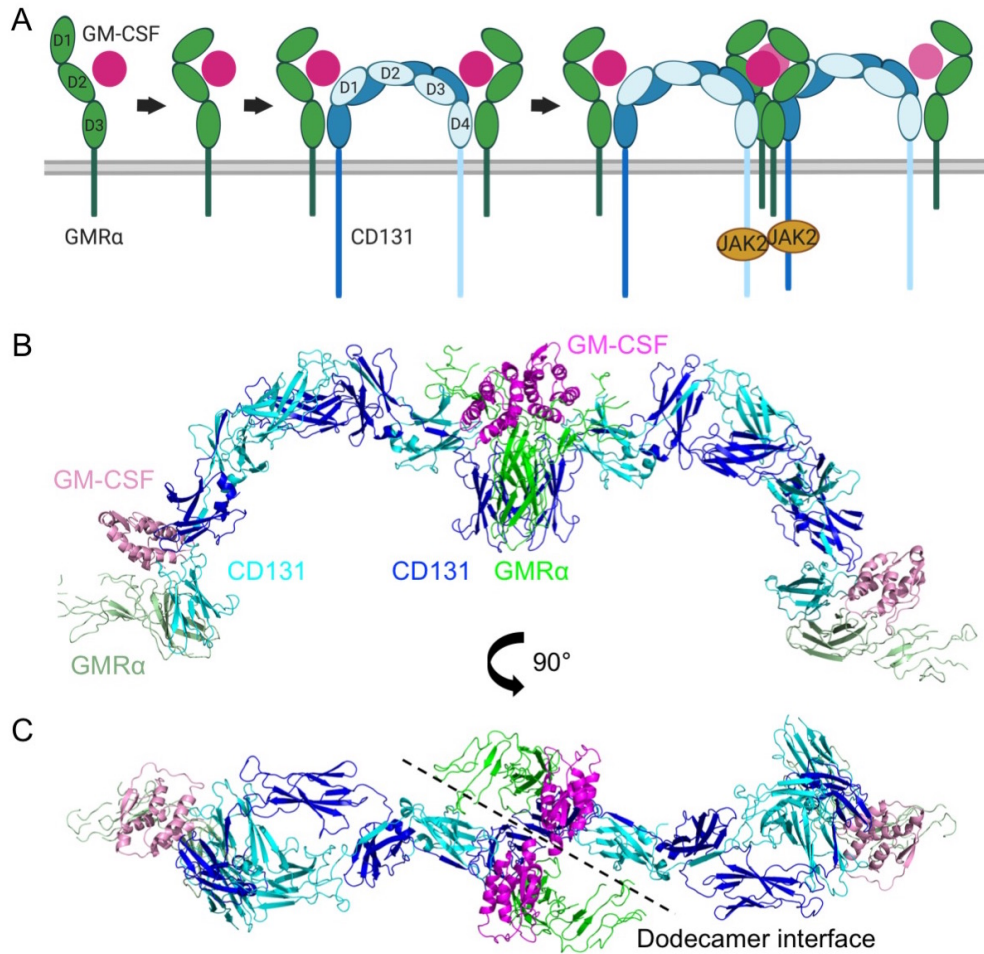
Figure 3.6: *GM-CSF signaling complex. (A) Mechanism of assembly of GM-CSF signaling complex. GM-CSF initially forms a low affinity binary complex with GMRα and then forms a high affinity complex with CD131. The complete signaling complex is a dodecamer comprising four molecules of GM-CSF, GMRα, and CD131. (B) Side and (C) top-down views of the GM-CSF–GMRα–CD131 dodecameric complex (PDB ID: 4NKQ). Some images were created with BioRender.com.*

EPO interacts with homodimeric EPO-R via two asymmetric binding interfaces. One side of EPO binds to EPO-R with strong affinity ($K_D \approx 1$ nM) while the other side binds with much weaker affinity ($K_D \approx 1$ μM) (Figure 3.2B) [16]. On the other hand, there is not enough structural and kinetics evidence to speculate the binding mode of EPO in complex with EPO-R and CD131. In this section, the structure of the EPO–EPO-R–CD131 heterocomplex was modeled based on the known structure of a signaling complex of another type 1 cytokine, GM-CSF, that requires CD131 for receptor activation and signaling and has a similar three-dimensional structure to EPO.

The structure of GM-CSF in complex with its receptor alpha subunit, GMRα, and CD131 has been extensively studied, and a set of crystal structures (PDB ID: 2GYS, 3CXE, 4NKQ, 4RS1) collectively provides information about the GM-CSF signaling complex. GM-CSF first forms a binary complex with a ligand-specific, low affinity GMRα with a K_D of about 2–10 nM. Then CD131, which lacks affinity for GM-CSF alone, binds to the binary complex and forms a high affinity ternary complex with a K_D of about 50–100 pM [73, 102]. During the transition to the ternary state, GM-CSF adopts a new conformation by tilting, such that the appropriate side chains are exposed to CD131 in the right orientations for optimal interactions [74]. Because CD131 exists as an antiparallel interlocked homodimer, each CD131 dimer forms two sets of ternary complexes, generating a hexamer. This hexamer then assembles into a complete dodecameric signaling complex, such that the cytoplasmic tails of two CD131s are close enough to induce JAK2 dimerization and phosphorylation. In this structure, a single GM-CSF interacts with GMRα and the domains 1 (D1) and 4 (D4) of each of the two CD131s in a homodimer, and this minimal structure resembles that of EPO–(EPO-R)$_2$ (Figure 3.6).

Extrapolating from the known structures of the EPO–(EPO-R)$_2$ complex and the GM-CSF–GMRα–CD131 complex, I hypothesized four possible ways that EPO could form a functional signaling complex with EPO-R and CD131.

1. EPO may bind to EPO-R via the strong side and to CD131 via the weak side. This is kinetically similar to the mechanism in which GM-CSF binds to GMRα and CD131. The ligand binds to the ligand-specific receptor subunit first, and then the binding of CD131 locks the complex into a high-affinity signaling conformation.

2. EPO may bind to EPO-R via the weak side and to CD131 via the strong side. This scenario is kinetically more challenging because CD131 has to compete with EPO-R for the strong side of EPO. However, this hypothesis could be true if EPO behaves differently from other type 1 cytokines and exhibits greater affinity for CD131 than for EPO-R on the strong side, or if the binding of CD131 to the strong side of EPO is driven by the relative abundance of CD131 on the cell surface. Alternatively, EPO-R and CD131 might exist as a pre-formed receptor complex so that CD131 can bind to the strong side of EPO preemptively [103].

3. EPO may bind to EPO-R via the strong side and to CD131 via regions of EPO other than the

strong and weak sides, such as the AB loop and the N- and C-termini. It has been shown that peptide fragments derived from helix B and the AB loop of EPO retained tissue-protective effects but not erythropoietic effects, demonstrating that these regions are sufficient to activate receptors for tissue protection [81, 82].

4. CD131 may not directly interact with EPO but may stabilize a receptor conformation that is amenable to tissue-protective signaling.

To learn which configurations may be structurally plausible, 10 structural alignment models of EPO–EPO-R–CD131 were created by aligning the structure of EPO–(EPO-R)$_2$ (PDB ID: 1EER) to that of GM-CSF–GMRα–CD131 (PDB ID: 4NKQ) in PyMOL. In most of these models, side chains of EPO and CD131 slightly overlapped possibly due to the lack of an energy minimization step. Models that showed steric clashes at the dodecameric interface were not excluded from further analyses, as structural models of the IL-5 signaling complex showed that structures equivalent to dodecameric complexes of GM-CSF or IL-3 may not be the only possible signaling unit for type 1 cytokines. Alignment of (IL-5)$_2$-IL-5Rα to the GM-CSF–GMRα–CD131 complex revealed that the formation of such a high-order complex may be sterically impossible for some type 1 cytokines [104].

The first two models were generated by aligning EPO–(EPO-R)$_2$ to GM-CSF in complex with GMRα as well as D1 and D4 of CD131, as these two structures exhibit high structural similarity. Interestingly, in this process, the use of two different algorithms resulted in two distinct modes of formation of the EPO–EPO-R–CD131 heterocomplex (models 1 and 2). CD131 interfaced with the weak and strong sides of EPO in models 1 and 2, respectively (Table 3.2 and Figure B.4A,B). When the complete dodecameric complex of GM-CSF–GMRα–CD131 was used instead, CD131 was positioned at the weak side of EPO (model 3) (Table 3.2 and Figure B.4C,D). The alignment of EPO to GM-CSF (model 4) also resulted in the binding of CD131 on the weak side of EPO (Table 3.2 and Figure B.4E,F). The algorithm used for generating models 1 and 3 (Align) takes into account the sequence identity between proteins that are being aligned whereas the algorithms used for models 2 and 4 (Super and CE, respectively) do not. This suggests that the weak side of EPO has higher sequential and structural similarity to the part of GM-CSF that binds CD131 compared to the strong side of EPO but that the structures of the whole receptor–ligand

48

Table 3.2: *Structural alignment approaches to generate structural models of EPO–EPO-R–CD131 heterocomplexes. EPO-R(s) and EPO-R(w) indicate EPO-R interacting with EPO via the strong and weak sides, respectively.*

No	Align X (PDB ID: 1EER)	To Y (PDB ID: 4NKQ)	PyMOL alignment algorithm	RMSD (Å)	EPO region interacting with CD131	Notes
1	EPO–(EPO-R)$_2$	GM-CSF–GMRα–CD131 D1/D4	Align	10.50	Weak side	
2			Super	14.58	Strong side	
3	EPO–(EPO-R)$_2$	GM-CSF–GMRα–CD131	Align	16.83	Weak side	
4	EPO	GM-CSF	CE	4.15	Weak side	Slight clash between EPOs at dodecamer interface
5	EPO-R(s)	GMRα	Super	4.00	Weak side	No contact between EPO and CD131
6	EPO-R(w)	CD131	CE	3.20	Weak side	Complete overlap between two EPOs as well as between EPO-R and CD131 at dodecamer interface
7	EPO–EPO-R(s)	GM-CSF–GMRα	Super	8.95	Weak side	Clash between EPOs at dodecamer interface
8	EPO-R(s)	CD131	CE	3.84	Strong side	
9	EPO-R(w)	GMRα	Super	3.40	Strong side	No contact between EPO and CD131
10	EPO–EPO-R(w)	GM-CSF–GMRα	Super	12.81	Strong side N-/C-termini	

complexes resemble better when CD131 faces the strong side of EPO.

For the configuration in which EPO's weak side interacts with CD131, three additional structural models were generated (models 5–7) (Table 3.2). For the other configuration in which EPO's strong side binds to CD131, three structural models were generated in a similar manner (models 8–10) (Table 3.2). The alignment of EPO-R to GMRα (models 5 and 9) did not result in any meaningful contact between EPO and CD131. The alignment models that simulated the binding of CD131 on the weak side of EPO (models 6 and 7) resulted in steric clashes between two EPOs of the neighboring hexamers, disrupting the dodecamer formation (Table 3.2 and Figure B.5). The formation of a dodecameric complex is essential for GM-CSF and IL-3 signaling but EPO might not necessarily rely on the dodecamer formation for successful signaling, as shown by the IL-5 complex. The other two models showing the binding of CD131 on the strong side of EPO (models 8 and 10) resulted in meaningful contacts between EPO and receptors without any significant steric clashes (Table 3.2 and Figure B.6).

EPO residues that might be important for receptor–ligand interaction and receptor activation were garnered from those within 4 Å of CD131 (Table 3.3). Models 1, 3, 4, 6, and 7, in which

Table 3.3: *EPO residues that are within 4 Å of CD131 domains 1 and 4 in structural alignment models. Those that make polar contacts with CD131 are underlined. Those that are mutated in this chapter are shown in bold.*

No.	Align X (PDB ID: 1EER)	To Y (PDB ID: 4NKQ)	PyMOL alignment algorithm	EPO residues within 4 Å of CD131
1	EPO–(EPO-R)$_2$	GM-CSF–GMRα–CD131 D1/D4	Align	Leu5, Ile6, Cys7, Asp8, Ser9, **Arg10**, Val11, Leu12, <u>Glu13</u>, **Arg14**, **Tyr15**, Leu16, Leu17, <u>Glu18</u>, Ala19, Glu21, Gln78, Leu93, His94, Val95, Asp96, Lys97, Ala98, Val99, <u>Ser100</u>, Gly101, Leu102, <u>**Arg103**</u>, **Ser104**, Leu105, Thr106, **Leu108**, Tyr145, Lys154
2			Super	Asp8, Ser9, <u>**Arg10**</u>, Val11, Leu12, Glu13, **Arg14**, Leu16, Leu17, <u>Lys20</u>, <u>Glu23</u>, Lys24, Thr27, Val46, Phe48, <u>Ser146</u>, <u>Asn147</u>, Leu149, <u>**Arg150**</u>, Gly151, Lys154, Leu155
3	EPO–(EPO-R)$_2$	GM-CSF–GMRα–CD131	Align	Asp8, Ser9, <u>**Arg10**</u>, Val11, Leu12, <u>Glu13</u>, **Arg14**, **Tyr15**, Leu16, Leu17, Glu18, Ala19, <u>Glu21</u>, Ala22, <u>Gln78</u>, Val82, Gln92, Asp96, <u>Lys97</u>, Ala98, Ser100, Gly101, <u>Tyr145</u>
4	EPO	GM-CSF	CE	<u>Asp8</u>, **Arg10**, Val11, <u>Glu13</u>, **Arg14**, Leu17, Glu18, Leu93, <u>Lys97</u>, <u>Ser100</u>, **Ser104**
5	EPO-R(s)	GMRα	Super	None
6	EPO-R(w)	CD131	CE	Leu5, <u>Asp8</u>, **Arg10**, Val11, Leu12, <u>**Arg14**</u>, **Tyr15**, Leu17, Glu18, Asp96, Lys97, Ser100, Gly101, **Arg103**, **Ser104**, Thr107
7	EPO–EPO-R(s)	GM-CSF–GMRα	Super	Leu5, Ile6, Asp8, <u>Ser9</u>, **Arg10**, Val11, Leu12, <u>Glu13</u>, <u>**Arg14**</u>, **Tyr15**, Leu17, **Leu108**
8	EPO-R(s)	CD131	CE	**Arg10**, Glu13, Leu16, Lys20, Thr44, **Lys45**, Val46, <u>Asn47</u>, Phe48, Tyr49, Lys52, <u>**Arg150**</u>, Lys154, Leu155, Arg162
9	EPO-R(w)	GMRα	Super	None
10	EPO–EPO-R(w)	GM-CSF–GMRα	Super	Ala1, Pro2, Pro3, Arg4, Cys7, <u>Asp8</u>, <u>Ser9</u>, **Arg10**, Tyr49, <u>Lys52</u>, Arg53, Met54, Lys154, Leu155, Thr157, Gly158, Glu159, Ala160, Cys161, <u>Arg162</u>, <u>Thr163</u>, Gly164, Asp165, Arg166

CD131 binds at the weak side of EPO, showed that EPO residues in helices A and C could interact with the D1 and D4 domains of CD131. Arg10 and Arg14 of EPO were located in proximity to CD131 in all of these models. Tyr15, Arg103, Ser104, and Leu108 of EPO were found in some of these models (Table 3.3 and Figure B.7A,C–F). These results are consistent with the protein docking models generated by Shing *et al.*, 2018. In these docking models, EPO interacts with CD131 via the weak side, and Arg10, Arg14, and Tyr15 of EPO are predicted to be crucial for EPO–CD131 interaction [76]. On the other hand, models 2, 8, and 10, in which CD131 binds at the strong side of EPO, had contact residues in helices A and D as well as in the AB loop. Arg10 was found to be in close contacts with CD131 residues in all three models. More interestingly, Lys45 and Arg150 of EPO, which abolished the tissue-protective effect of EPO upon mutation,

were within 4 Å of CD131 in models 2 and 8 (Table 3.3 and Figure B.7B,G). In model 10, EPO residues at the N- and C-termini pointed at CD131 but those lining the major helices and loops were not directly in contact with CD131, supporting the possibility that parts other than the strong and weak sides of EPO may interact with CD131 (Table 3.3 and Figure B.7H). Structural models based on the GM-CSF receptor complex showed that all four hypotheses for the formation of EPO–EPO-R–CD131 are plausible and that EPO residues that are close to CD131 are in a good agreement to those that were selected for mutagenesis in this studies as well as those resulted from Shing *et al.*, 2018's docking studies [76].

Previous mutagenesis studies on the GM-CSF signaling complex show that mutations that abolish bioactivity of the ligand are the ones that dramatically disrupt the binding of CD131, not the receptor alpha subunit. For instance, GMRα(K195D) completely lost GM-CSF binding but reduced GM-CSF activity only by about 10-fold in vitro. Similarly, when the affinity of GM-CSF(D112K) for GMRα was reduced by about 230-fold, the in vitro cell-proliferative activity of the same mutant was reduced only by about 8-fold [74]. These results imply that even total disruption of ligand binding to the receptor alpha subunit may not necessarily block bioactivity of the ligand, and highlight the importance of affinity conversion by CD131. This is consistent with previous observations that mutations that disrupt the GM-CSF–CD131 interaction completely eliminated cellular responses to GM-CSF [105, 106]. If the EPO–EPO-R–CD131 heterocomplex follows a similar assembly mechanism as the GM-CSF signaling complex, then one can deduce that residues that remove tissue-protective effects but not erythropoietic effects upon mutation are likely to be essential for CD131 interaction. None of the EPO mutants tested in this study ablated tissue-protective activity while still maintaining erythropoietic activity independently of IH4 fusion. Thus, further mutagenesis studies are needed to search for residues that are important for CD131 interaction, if EPO and CD131 indeed form direct contacts.

Next, experimental data were examined under the four hypothetical scenarios proposed above. In in vitro cell-based assays, mutating Lys45 or Arg150 resulted in the loss of both erythropoietic and tissue protective activity (Figure 3.3 and Figure 3.5D,E). This behavior of the Lys45Asp (K45D) and Arg150Ala (R150A) mutants does not rule out any hypothesis. Erythroid cell proliferation assay data imply that these two mutants interfere with EPO-R binding at the strong side. Hence, their loss of tissue-protective effects may also be caused by the disruption of

51

EPO interaction with EPO-R. Hypotheses 1, 3, and 4, in which the strong side of EPO may interact with EPO-R, fit with this set of data. Hypothesis 2, in which the strong side of EPO may interact with CD131, cannot be rejected by these data either because tissue protective activity of R150A and K45D may be significantly reduced if these mutants hinder with the binding of CD131.

In contrary, the activity of the Ser104Ile (S104I) and Leu108Ala (L108A) mutants ruled out the possibility of CD131 binding at the strong side of EPO (hypothesis 2). Erythroid cell proliferation assay data show that these mutants completely ablate EPO-R homodimer signaling by disrupting the receptor interaction at the weak side of EPO (Figure 3.3) [87]. This implies that the binding of EPO to EPO-R and CD131 via the weak and strong sides, respectively, should also be disrupted by these mutants. However, these mutants were able to protect neuronal cells from cobalt-induced death in cell-based assays (Figure 3.5C,H,I), providing a counterevidence for hypothesis 2. These data do not reject the other three hypotheses, as these mutations may not reduce CD131 binding affinity strongly enough to affect signaling (hypothesis 1), or may not directly interact with CD131 at all (hypotheses 3 and 4).

Lastly, erythropoietic and tissue-protective activity of darbepoetin (Figure 3.3 and Figure 3.5B) rules out the possibility of EPO interaction with CD131 via the pointy end of the prolate-spheroid-shaped EPO that is opposite to the N- and C-termini (that is, regions near the AB and BC loops). Darbepoetin has two additional *N*-linked glycosylation sites (Asn30 and Asn88) engineered into wild-type EPO [107]. The added sugar moieties are located at a region of EPO that does not directly bind to EPO-R but are bulky and negatively charged so that they cause steric and electrostatic hindrance with receptors, as evidenced by slightly reduced erythroid cell proliferative activity compared to wild-type EPO (Figure 3.3A). Despite these bulky changes, darbepoetin was able to promote tissue protection in vitro, suggesting that CD131 likely does not form a direct contact with EPO via the region with extra glycosylation. These data partially but not entirely rebut the hypothesis that CD131 may bind EPO via regions other than the strong and weak side (hypothesis 3).

Structural alignment models and cell-based experiments show that three out of four proposed hypotheses could explain the potential binding mode of the EPO–EPO-R–CD131 complex. The second hypothesis that EPO may interact with EPO-R and CD131 via the weak and strong sides, respectively, is kinetically less favorable than the other hypothesized configurations, and is

eventually rejected by the weak-side EPO mutants that failed to stimulate erythroid cell proliferation but exhibited tissue-protective activity. Although these analyses help visualize potential three-dimensional structures of the EPO–EPO-R–CD131 heterocomplex and estimate the regions of EPO that might be involved in receptor activation for tissue protection, more rigorous structural and mutagenesis studies need to be performed to validate or disprove the proposed hypotheses and structural alignment models, and to broaden our understanding of mechanisms of the pleiotropic actions of EPO.

3.4   Conclusions

Attachment of different EPO mutants to an antibody element enabled us to mix and match different flavors of EPO action for various therapeutic applications. For instance, a form of EPO that stimulates both RBC production and tissue protection without pro-thrombotic side effects would be useful for the treatment of high-altitude-related illnesses, as well as for general performance enhancement and prophylactic protection from chemical and physical damage in the military. A version of EPO that stimulates only tissue protection without RBC production and pro-thrombotic side effects would be useful for the treatment of TBI and neurodegenerative diseases as well as surgery enhancement. In addition to a broader spectrum of therapeutic applications, incorporating a mutation that significantly reduces side effects and maintains the desired effect would increase the maximum allowed dose of EPO in clinical studies, thereby allowing new studies to be conducted at more clinically relevant doses. Currently, the maximum allowed dose of EPO is limited by its known effects on anemia in patients with chronic kidney diseases [108]. The effective dose of EPO for protection of neuronal cells, however, is predicted to be much higher than that for the treatment of anemia because only <1% of administered EPO is able to cross the BBB, drastically reducing the local concentration of EPO available for target cells [109]. If doses are limited by the value set for anemic patients, then EPO is likely to fail clinical trials for neuroprotection because doses are insufficient to reach an effective local concentration in the target tissue and not necessarily because the drug itself is not effective. Engineering strategies described in this chapter enable us to harness different sets of beneficial effects of EPO according to different medical needs, and to improve the chances of successful clinical trials by allowing the

administration of more clinically relevant doses.

3.5   Materials and Methods

Materials and Methods for cell culture of 293-F, CHO-S and CHO DG44 cell lines, DNA construction of EPO variants, protein expression and his-tag purification, in vitro measurement of erythropoietic activity by cell proliferation assays, and measurement of mouse reticulocytes and reticulated platelets by flow cytometry were described in Chapter 2.

Cell Culture. Human neuroblastoma cell line, SH-SY5Y, was obtained from ATCC. SH-SY5Y cells were cultured in 1:1 DMEM/F-12 with 10% FBS at 37°C in 5% CO_2.

Neurorotection Assay. SH-SY5Y cells were seeded in a 96-well plate at $4.0x10^4$ cells per well in 80 µL of 1:1 DMEM/F-12 with 10% FBS, and were let adhere overnight in the incubator (37°C, 5% CO_2). Cells were pre-treated with varying concentrations of purified proteins for 24 hr at 37°C in 5% CO_2. Then 100 µM of cobalt chloride ($CoCl_2$), a hypoxia-mimicking agent, was added to cells. After 24 hr, cell viability was measured by CellTiter 96$^®$ AQueous One Solution Cell Proliferation Assay (Promega). Absorbance at 490 nm was read on a BioTek Synergy Neo HTS microplate reader. Reported data represent mean $\pm$ S.E.M of two or three replicates.

Structural Alignment. The crystal structures of EPO–(EPO-R)$_2$ (PDB ID: 1EER) and GM-CSF–GMRα–CD131 (PDB ID: 4NKQ) were obtained from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB; rcsb.org). The structural alignments of these two receptor–ligand complexes were performed and RMSD values were calculated using three algorithms, Align, Super, and CE, in PyMOL. Super and CE were preferred to Align, as the sequence identities between EPO and GM-CSF and between EPO-R and GMRα or CD131 were lower than 30%. Super and CE were excluded from analyses of some alignment models due to the following limitations. Super had a tendency to align EPO-R with CD131 homodimer interchain region, orienting domain 2 of EPO-R parallel to the membrane. CE aligned solely based on protein domains with most structural similarity, ignoring the rest of the complex being aligned. Residues that are within 4 Å of interacting receptor or ligand were selected using PyMOL.

Chapter 4

Resurfacing Erythropoietin to Improve Expression and Reduce Immune Responses that Lead to Pure Red Cell Aplasia

4.1 Abstract

Major challenges in developing protein therapeutics are poor expression of recombinant proteins in host cell systems and potentials for immune responses to the drug. Recombinant human erythropoietin (EPO) has been used to treat anemia in chronic renal failure patients and cancer patients receiving chemotherapy. Although EPO is naturally found in the body, synthetic drugs can be recognized as foreign by the patient's immune system due to the presence of misfolded proteins, aggregates, or contaminants acquired during production, formulation, or storage. In this case, an immune response to the administered EPO may result in anti-drug antibodies, which not only sequester away drug action but also neutralize the patient's own endogenous EPO, completely removing the stimulus required for red blood cell production in the bone marrow. In this chapter, the surface residues of EPO were mutated such that anti-drug antibodies, if generated, would not cross-react with the endogenous hormone. Additionally, these

mutations were introduced such that they would stabilize protein folding for better expression. Protein sequence homology among mammals and structural features of EPO were examined to determine EPO residues to modify. In this process, 12.0–27.1% of the total amino acid residues and 12.3–29.5% of the solvent-exposed surface residues were mutated from wild-type EPO. Surprisingly, highly resurfaced EPO variants were expressed at higher levels compared to the least modified variant. Nine of ten variants maintained their ability to stimulate cell proliferation in vitro. When these variants were used to build Targeted EPO fusion proteins, six of them induced cell proliferation in a glycophorin A-dependent manner in vitro. Their potential for generating cross-reactive antibodies will be further evaluated by measuring the binding of these fusion proteins to polyclonal antibodies, or anti-sera, for human wild-type EPO. The current work shows that mutating up to 27% of the total amino acid residues does not jeopardize EPO activity, and could increase protein expression, if the modifications are designed in a rational manner. It also demonstrates that rational engineering strategies based on protein sequence and structure can inform smarter drug design that could mitigate immunological side effects.

4.2   Introduction

4.2.1   Expression of Recombinant Cytokines and Hormones

Cytokines and hormones are classes of molecules that regulate a broad spectrum of physiological events through cell signaling. They are naturally produced by various types of cells in our body to respond to changes in the internal environment of the body and maintain homeostasis. They bind to receptors on the target cell surface, force receptors into active conformations, and activate intracellular signaling pathways to induce changes in the cellular functions. These signaling events are turned off when the activated receptor–ligand complexes are internalized by clathrin-mediated endocytosis and processed for degradation or recycling. This receptor-mediated endocytosis ensures that a cell does not receive too much signaling, which could lead to aberrant or uncontrollable cellular behavior.

The implication of this mechanism is that cytokines and hormones are designed by the nature to unfold easily such that they are degradable upon signaling and receptor-mediated endocytosis. This feature is favorable in a natural setting in which balance between constant

production and degradation of cytokines and hormones helps maintain homeostasis of the body. However, in the context of protein therapeutic production, this feature can make protein expression difficult. Interferon is an example of a cytokine whose recombinant form is expressed very poorly in a host cell system (yields low mg/L, compared to monoclonal antibodies that yield about 1–10 g/L; data not shown) [110]. Poor protein expression not only significantly increases the cost of production of a protein drug but also limits the development of a new drug derived from a cytokine or a hormone.

4.2.2 Immunogenicity of Protein Therapeutics

The human body is intricately designed to respond to and fight against foreign molecules that look unfamiliar and dangerous. The innate immune system provides first-hand responses against potentially harmful intruders, such as bacteria and viruses, and alerts other immune cells. The adaptive immune system forms memory of previously exposed antigens, and initiates faster defense against the same antigens at the next encounter. These mechanisms may guard our body from infections but they can be problematic when they act against therapeutic proteins. Synthetic hormones and cytokines, for instance, are generated in a laboratory by expression of recombinant DNA in a production cell line, and are used to replace or supplement their endogenous counterparts in patients who lack or have low levels of them. Although they are designed to be exact copies of proteins that are naturally found in the body, they can adapt a slightly different structure, expose an otherwise hidden epitope, aggregate, or even acquire contaminants during production and storage [111, 112]. Therefore, the ability of protein therapeutics to instigate an immune response, or immunogenicity, is a great concern in drug development.

Immunogenicity of protein therapeutics arises via a similar mechanism to the classical immune responses to pathogens or vaccines. Foreign proteins are recognized by immune cells, and lead to B cell activation and antibody secretion. This process involves many steps and can occur in a T cell-independent or a T cell-dependent manner. In T cell-independent B cell activation (Figure 4.1A), B cells are activated by multivalent molecules, such as polymers, carbohydrates, and aggregates, which cluster and activate multiple B cell receptors (BCRs) simultaneously, providing

strong signals needed for the generation of antibodies. However, this process results only in short-lived plasma B cells, which produce weak affinity IgM [111, 112, 113].
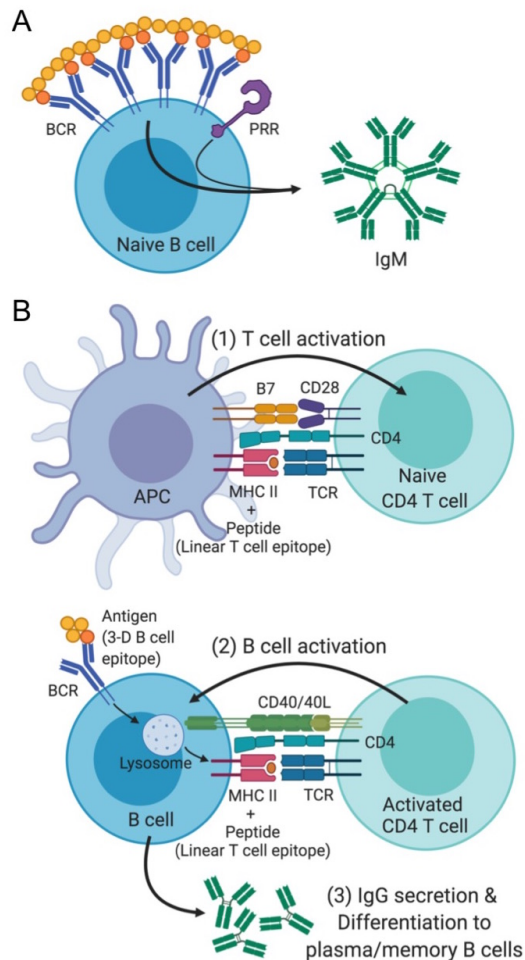
Figure 4.1: *Mechanisms of antibody production by lymphocytes. (A) T cell-independent B cell activation. B cell receptors (BCRs) recognize three-dimensional motifs of polymeric repeats, aggregates, or polysaccharides. Pattern recognition receptors (PRRs) sense danger signals. Clustering and simultaneous activation of multiple BCRs, together with the danger signals, triggers the production of weak affinity IgM antibodies. (B) T cell-dependent B cell activation. An antigen presenting cell (APC) presents a linear peptide fragment from an antigen on MHC II and relays a danger signal by upregulating B7. A naïve CD4 T cell is activated when it recognizes the antigenic peptide via a T cell receptor (TCR), and receives the danger signal via CD28. The activated CD4 T cell activates the B cell that presents the same peptide. The activated B cell proliferates, secretes IgG antibodies, and differentiates into long-lived plasma cells or memory B cells. Images were created with BioRender.com.*

Most antibodies against protein therapeutics develop through T cell-dependent B cell activation (Figure 4.1B). In this process, BCR binds to an antigen with a specific three-dimensional structure, whether it be multivalent or monovalent, and mediates the endocytosis of the receptor–antigen complex. The antigen then is degraded in the lysosome, and its protein components

are presented on MHC II as linear peptides of 8–15 amino acids. The antigen-presenting B cells then migrate to the T–B interface in the lymph node. In the T cell zone, naïve CD4 T cells are activated by antigen-presenting cells (APCs). APCs, such as macrophages and dendritic cells, routinely sample proteins and pathogens from their surroundings, digest them into linear 8- to 15-mer peptides, and present these peptides on MHC II. If pattern recognition receptors (PRRs), such as toll-like receptors (TLRs), are stimulated at the same time, then the APCs are brought closer to the T cell zone in the lymph node and co-stimulatory molecules ("danger signal") are upregulated on the cell surface. When the digested peptide on MHC II and the danger signal are recognized by a T cell receptor (TCR) and co-stimulatory receptor, respectively, the naïve CD4 T cells become activated and migrate to the T–B interface in the lymph node. There, the antigen-presenting B cells are activated by the T cells that were previously activated by the same peptide that the B cells are presenting. The activated B cells initiate antibody secretion, isotype switching, and somatic hypermutation, and further activate T cells to aid in the selection and survival of B cells that further differentiate into antibody-secreting plasma cells or memory B cells. This crosstalk between T and B cells results in a more robust and long-lasting adaptive immunity that is capable of faster release of high affinity antibodies [111, 112, 113].

After the initial administration of the drug that contains a foreign epitope, the set of immune responses described above takes place and weak affinity antibodies are secreted within about 4–7 days in humans. Upon sequential administrations of the same drug, the immune responses are much faster, taking about 1–3 days in humans to produce high affinity antibodies from the plasma cells and memory B cells that were formed during the first exposure [113]. The consequences of anti-drug antibodies could be as trivial as reduced drug efficacy but could also be very detrimental if anti-drug antibodies are cross-reactive with native proteins that share the same antigenic epitopes. Therefore, careful design and manufacturing of protein therapeutics are particularly important when multiple administrations are recommended for the desired clinical outcome in patients.

4.2.3 Erythropoiesis-Stimulating Agents and Pure Red Cell Aplasia

Erythropoiesis-stimulating agents (ESAs) are therapeutics that stimulate the production of red blood cells (RBCs), or erythropoiesis, and are used to correct anemia caused by chronic kidney diseases or chemotherapy. The first ESA, recombinant human erythropoietin (EPO), was approved and introduced to the market in 1989 in the United States. Since then, various forms of recombinant human EPO, such as epoetin alfa and epoetin beta, have been commercialized. Although recombinant EPO is made as an exact homolog of the endogenous counterpart, unavoidable differences in glycosylation patterns and protein folding pose concerns for immunogenicity. Such concern is especially important for ESAs in comparison to other protein therapeutics because immune responses against ESAs result in a sudden and severe form of anemia, pure red cell aplasia (PRCA).

PRCA is a type of anemia described by the cessation of RBC production due to a complete or near-absence of erythroid precursor cells in bone marrow that otherwise looks normal in terms of the number and morphology of other lineage cells. It can be caused by autoimmune diseases, genetic disorders, viral infections, and ESA therapy [114]. ESA-induced PRCA is very rare, with an exposure-adjusted incidence of 2–3 per million patient-years [115]. However, with occasional surges in cases of ESA-induced PRCA and increasing production of biosimilars worldwide, ESA-induced PRCA should be considered as a serious adverse effect that requires a non-trivial set of treatments, including blood transfusion and immunosuppressive therapy, which significantly reduce one's quality of life.

Between 1998 and 2004, the number of reported cases of antibody-induced PRCA in ESA-treated patients surged to 191, more than half of which were in France, Canada, Spain, and the United Kingdom. Most of these patients received ESAs marketed outside the United States (Eprex® (epoetin alfa), Johnson & Johnson; Neorecormon® (epoetin beta), Roche). Five out of the 191 patients received epoetin alfa marketed in the United States (Epogen® (epoetin alfa), Amgen) [116]. Analyses of serum samples from 13 patients with chronic kidney failure who developed PRCA after epoetin treatment revealed anti-EPO antibodies that inhibited erythroid colony formation by normal bone marrow cells, indicating that the condition is mediated by neutralizing antibodies against ESA that cross-react with endogenous EPO. This study also showed

that those neutralizing antibodies recognized conformational epitopes in the protein, not the glycan moiety of EPO [117]. However, how such antibodies were generated and why the number of cases suddenly increased between 1998 and 2004 are not fully understood. Many factors are speculated to contribute at least partially to the onset of immune responses. For instance, most patients received ESAs by subcutaneous administration, which exposes the drug to a high local concentration of immune cells in the skin for a longer duration of time due to slow resorption. Most cases were in patients that received Eprex® formulated without human serum albumin, which, when added to drug formulations, helps with protein stability and reduces aggregate formation. Mishandling of the drug during distribution and storage damages protein stability and induces aggregate formation. There is not any experimental evidence to validate these hypotheses, but stricter measures for appropriate handling, storage, and the route of administration lowered the incidence rate of ESA-induced PRCA by more than 80% worldwide, relative to the peak incidence in 2001 [115, 116].

Upon diagnosis of PRCA, cessation of ESA treatment and immunosuppressive therapy are required to reduce the quantity of anti-EPO antibodies in circulation and achieve hematologic recovery [114, 118, 119]. ESA-induced PRCA itself is not fatal, and most patients recover within a few months of immunosuppressive therapy. However, these patients become highly dependent on blood transfusion, requiring about one unit of RBC per week, and are thus at risk of transfusion-transmitted infections as well as iron overload [114]. Because anti-EPO antibodies are highly cross-reactive and neutralize all other commercially available ESAs as well as endogenous EPO, it is unclear if ESA therapy can be safely resumed in these patients. A follow-up study on patients with ESA-induced PRCA suggests that re-initiation of ESA therapy in patients who once had the condition could be considered if their anti-EPO antibodies are no longer detectable in blood. In this study, about 89% of the patients who had undetectable levels of anti-EPO antibodies after immunosuppressive therapy were responsive to ESA re-challenge. However, those that did not respond to ESA re-challenge remained heavily dependent on routine blood transfusion [118, 119].

4.2.4 Protein Engineering Strategies for De-Immunization

Anti-drug antibody formation is mediated by two mechanisms, T cell-independent and T cell-dependent B cell activation, as described in Chapter 4.2.1. The former takes place when a B cell recognizes a three-dimensional polyvalent epitope and produces low affinity IgM antibodies, without help from T cells (Figure 4.1A). The latter occurs when a CD4 T cell is activated by a linear T cell epitope presented on an APC along with a danger signal, and in turn, prompts a B cell carrying the same T cell epitope to produce high affinity IgG antibodies (Figure 4.1B). Negative immune responses to most protein therapeutics, including ESAs, arise from the T cell-dependent mechanism, as evidenced by the presence of high affinity IgG1 and IgG4 antibodies [120, 121].

Reducing the immunogenicity, or de-immunization, of protein therapeutics can be rationally performed via several protein engineering strategies. In particular, T cell-dependent anti-drug antibody formation can be mitigated in two ways. One approach is to modify linear peptide fragments that are predicted as potential T cell epitopes. In theory, this would block all formation of long-term, high affinity anti-drug antibody by suppressing the activation of T cells. T cell epitopes can be predicted in silico and mutated to reduce the fitness to the peptide binding groove of MHC II on APCs, thereby de-stabilizing the complex formation of MHC II and antigenic peptides and reducing the chances of T cell activation [111]. The immunodominant T cell epitopes of EPO have been previously identified and mutated to generate forms of EPO that elicited reduced immunogenicity in vitro [122]. However, it has been shown that almost half of normal donors have pre-existing CD4 T cells specific for human EPO in their blood, suggesting the presence of T cells that are ready to be activated when strong signals, such as poorly formulated drugs, are delivered to break the immune ignorance [123].

Another approach is to shield potentially foreign three-dimensional conformational epitopes from BCRs and PRRs. Unlike TCRs, which bind linear peptide fragments, BCRs bind three-dimensional structural elements that are not limited to proteins, including polymers, glycans, and lipids. Simultaneous relay of danger signals alerts B cells of foreignness and activates them for downstream responses. Modifying or introducing disulfide bond patterns can reduce the solvent exposure of the hydrophobic core, thereby reducing the visibility of epitopes that B cells might consider "foreign". The resultant increase in protein stability also reduces chances

of aggregate formation, thereby avoiding recognition by BCRs [124]. Engineering glycosylation patterns is another way to provide a "shielding" effect. Adding bulky glycans on the surface shields proteins and hides potentially foreign epitopes from B cells [125]. By preventing activation of B cells, production of plasma cells, memory B cells, and high affinity antibodies can be avoided even when activated CD4 T cells are present.

Although these strategies have shown some success in de-immunization of protein therapeutics in academic laboratories [122, 126], "rational" design of less immunogenic protein drugs is limited by our lack of understanding of the mechanisms of immunity against recombinant human proteins that are designed as copies of the "self" proteins, and is complicated by highly diverse immune systems between individuals. As a result, even proteins that are rationally engineered to reduce immunogenicity always carry a risk of triggering unexpected immune responses. The massive incidences of ESA-induced PRCA in the late 1990s and early 2000s are good examples of unforeseen immunogenicity that affected hundreds of people's quality of life. Although the protein itself was not engineered for de-immunization, one can see that immune responses can be triggered by many factors other than a protein's intrinsic properties, such as improper formulation and handling/storage methods [116]. Therefore, in addition to de-immunization of proteins, other engineering efforts are needed to ensure the safety of such drugs in case of unpredicted immune responses leading to adverse effects.

Such effects of anti-drug antibodies result when the antibodies cross-react with native proteins produced by patients and abolish the action of both the exogenous drugs and the endogenous proteins. Generation of such cross-reactive and neutralizing anti-drug antibodies can be avoided by resurfacing the protein to make it look different from the endogenous protein at a three-dimensional structural level. Most patients who developed ESA-induced PRCA had anti-EPO antibodies that recognized conformational epitopes on a folded protein structure, as opposed to linear epitopes on a denatured protein [117]. This suggests that cross-reactivity of antibodies may be avoided if the resurfaced proteins presented different conformational motifs from endogenous ones on the solvent-exposed surface. This concept was demonstrated by peginesatide (Omontys®, Affymax and Takeda), a synthetic peptide dimer linked by polyethylene glycol (PEG) that stimulates RBC production by binding and activating EPO-R. Because its amino acid sequence is completely different from that of EPO, antibodies against peginesatide did not

cross-react with EPO; conversely, those against EPO did not cross-react with peginesatide [127]. In clinical studies, peginesatide improved hemoglobin levels in patients with EPO-induced PRCA [128]. However, this drug was recalled from the market in 2013 after a clinical study showed that it had worse cardiovascular side effects and a higher death rate than darbepoetin [129]. The promise and downfall of peginesatide highlights that the combination of resurfaced EPO and the Targeted EPO chimeric activator described in Chapter 2 would result in a safer second-generation ESA. Targeted resurfaced EPO would carry a lower risk for pro-thrombotic and cardiovascular side effects. It would also enable continual ESA treatment with other agents in case PRCA is induced, ablating the need for routine blood transfusions and immunosuppressive therapies that significantly decrease quality of life.

4.3 Results and Discussion

The present work demonstrates that protein engineering strategies can improve the expression of a protein drug by introducing stabilizing mutations as well as the safety of a drug by preventing the formation of cross-reactive and neutralizing antibodies that suppress the activities of both the drug itself and the native protein produced by the patient's body. This work was motivated by the previous observation that the Targeted EPO molecule, IH4-5-EPO(R150A), described in Chapter 2, triggered PRCA in mice. In this experiment, mice received intraperitoneal (i.p.) injections of darbepoetin, IH4-5-EPO(R150A), or vehicle alone twice a week for up to four weeks. 50 μg of bovine serum albumin (BSA) was co-injected as an adjuvant in the darbepoetin and IH4-5-EPO(R150A) groups [130]. A small amount of blood was drawn by tail-nick on indicated days to measure reticulocyte and reticulated platelet counts (Figure C.1A). Mice treated with IH4-5-EPO(R150A) reached depletion of reticulocytes between Days 11 and 25, exhibiting a drastic decrease from 11.17% to −3.70% relative to Day 0 (Figure C.1B). All mice stopped receiving injections after a sign of drug-induced PRCA was noted on Day 25, and were monitored for an additional two weeks (Figure C.1). Mice in the IH4-5-EPO(R150A) group did not recover from anemia within the two weeks after cessation of dosing, during which their reticulocyte counts were consistently below baseline (Figure C.1B). By the end of the two week period, they appeared to be pale and showed piloerection (data not shown). In contrast, darbepoetin did not induce PRCA

over the course of this study. Mice treated with darbepoetin maintained elevated reticulocyte counts until Day 32, a week after injections were stopped, and then showed a dramatic drop on Day 38 (-4.31%); this normally occurs about 10 days after the removal of an exogenous supply of EPO due to a lack of high EPO signal and the exhaustion of the erythropoiesis machinery (Figure C.1B).

These results suggest that IH4-5-EPO(R150A) induced the formation of anti-drug antibodies, which also cross-reacted with and neutralized the endogenous EPO made by the mice. There are many potential causes for the immunogenicity of the IH4-5-EPO(R150A) used in this study: unexpected contaminants could have been introduced during the production, mishandling or bad formulation could have increased aggregate formation, and the intrinsic difference between human and mouse EPO sequences or structures could have led to immune responses. Despite having about 80% sequence homology, human and mouse EPOs differ in disulfide bond patterns, and human EPO contains an O-linked glycosylation site that is not present in mouse EPO (Figure C.2). Because darbepoetin, a hyperglycosylated form of human EPO, did not trigger immune responses in mice, it is more likely that extrinsic factors such as handling and formulation were responsible for the immunogenicity of Targeted EPO, and that the protein itself would not be immunogenic, if prepared and formulated more carefully following the industry-level practices. Likewise, any commercial protein therapeutics are at a risk of generating anti-drug antibodies and inducing adverse consequences. This experiment inspired us to use EPO as a model to design protein therapeutics that are as equally functional as their wild-type counterparts but do not generate neutralizing antibodies that cross-react with the endogenous proteins in case of any immune response to the drug.

4.3.1   Rationale for Resurfacing Mutations

The purpose of resurfacing EPO was to present surface epitopes that were different from the endogenous human wild-type EPO, such that if the exogenous EPO drug causes an immune response, antibodies against it will not cross-react and neutralize the endogenous EPO. Additionally, the implicit goal was to improve protein folding and expression of EPO. Therefore, it was essential that a good fraction of the surface residues be mutated and that these mutations

should not interfere with the expression, folding, and therapeutic activity of the engineered protein. I focused on three characteristics of EPO to ensure that the above requisites were satisfied: protein sequence homology among mammals, disulfide bond patterns, and glycosylation patterns.
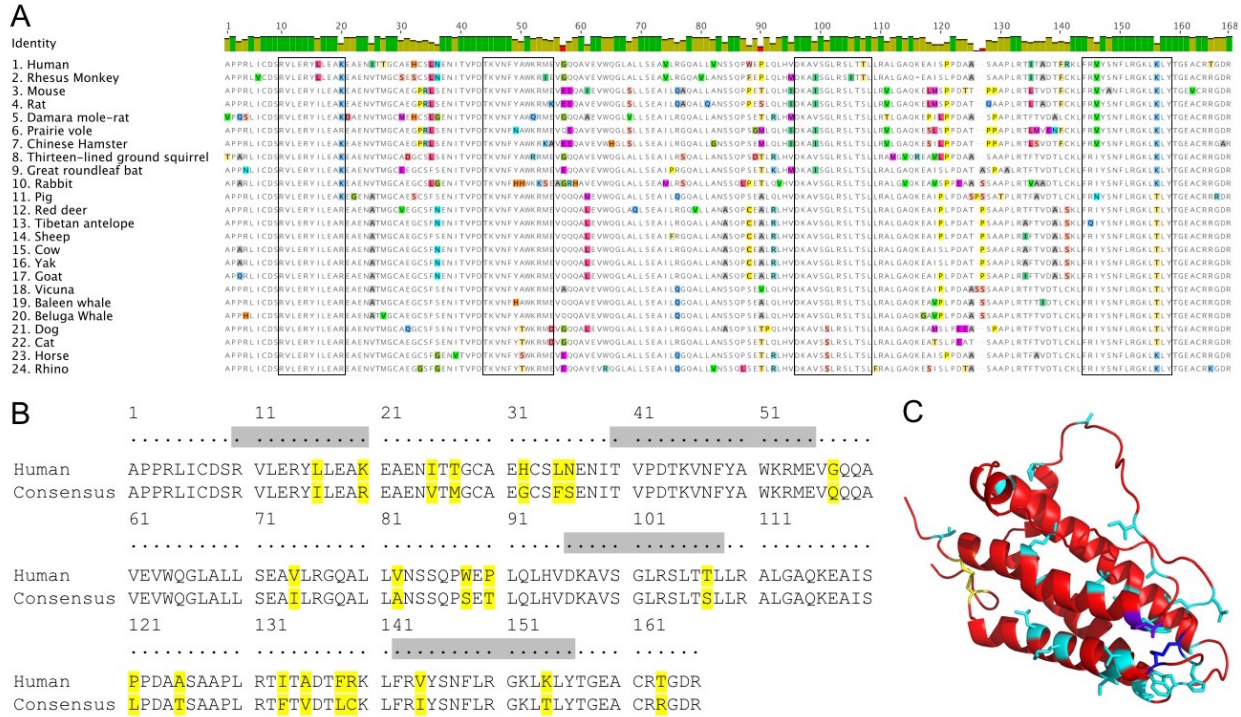
Figure 4.2: *Protein sequence and structural alignment of EPO from 24 mammals. (A) EPO protein sequences from all 24 mammals, including human, were aligned. (B) Human EPO was aligned with the consensus of 23 other mammalian EPO sequences. They show 86.8% sequence homology. In these sequence alignments, highlighted amino acids indicate differences between a given sequence and the consensus (A), or between the two aligned sequences (B). Boxed or gray regions are important for EPO–EPO-R interaction and activity. Sequences can be viewed in Appendix C. (C) Structure of human EPO (PDB ID: 1EER). Residues that are different from the non-human mammalian consensus are shown as cyan sticks. Disulfide bond (Cys7–Cys161) common to human and non-human consensus is shown in yellow. Disulfide bond (Cys29–Cys33) found in human but not in consensus is shown in blue. Residue at position 139 (Arg139), shown in purple, is a cysteine in consensus and likely forms an alternative disulfide bond with Cys29.*

First, EPO protein sequences from 24 mammals were aligned, and highly conserved and non-conserved regions were identified (Figure 4.2). EPO is a highly conserved protein with about 80% protein sequence homology among mammals, and human EPO functions in mice, rats, and dogs [124, 131]. Regions that are important for receptor–ligand interaction and activity, such as helix A (aa 10–20), the AB loop (aa 44–55), helix C (aa 96–108), and helix D (aa 142–156), are highly conserved among mammals (Figure 4.2) [95]. Regions that are composed of flexible loops or facing away from the receptors, such as the CD loop (aa 111–138) and the N-/C-termini, are the

66

most variable (Figure 4.2). Of particular interest were residues that are highly conserved among mammals but not in the human sequence. For instance, the position 27 of EPO is threonine in humans but methionine in 22 out of 23 other non-human mammals (Figure 4.2A). Likewise, the AB loop of human EPO contains three residues, His32, Leu35, and Asn36, which differ from the consensus sequence among the 23 non-human mammals (Figure 4.2B). Reverting the human sequence to the consensus at these positions would maintain protein folding and activity because highly conserved residues tend to be evolutionarily stable and beneficial for EPO function. In the regions where there is low sequence homology, reversion to the consensus is not a viable strategy. For instance, the CD loop is a long, flexible region of the protein that does not directly interact with the receptors, and is the least conserved in the pairwise alignment among mammals. It is generally rich in proline and small hydrophobic residues such as alanine, isoleucine, and leucine (Figure 4.2A). To introduce more variation in this region, some residues were mutated away from the human sequence or new residues were inserted to change the linear sequence and the peptide bond angles (Figure 4.3).

Second, disulfide bond patterns were modified to improve protein folding, stability, and activity. EPO has four to five cysteine residues in mammals. All mammals have conserved Cys7, Cys29, and Cys161 residues as well as cysteines at one to two of three other positions (33, 88, and 139). Human and rodent EPOs have Cys33 and Cys139, respectively. EPOs from ungulates, such as cows and sheep, have Cys33 and Cys88. Most other mammalian EPOs have Cys33 and Cys139 [124]. Human EPO with rearranged disulfide bond patterns (Cys29–Cys33 to Cys29–Cys88) was shown to express better with less aggregate formation, to be more stable to removal of *N*-linked oligosaccharides, and to have improved pharmacokinetics and pharmacodynamics in mice [124]. It was hypothesized that the Cys29–Cys88 bond would allow the most water-tight, energetically stable structure by covalently linking the AB and BC loops, which would otherwise be held together by hydrophobic contacts that are not shielded from the solvent [124]. The rearranged disulfide bonds are also expected to hide potentially foreign-looking epitopes by reducing the exposure of the hydrophobic core of the alpha helical bundle to immune cells. The same mutations were introduced in some of the resurfaced EPO variants to rearrange disulfide bonds and improve protein folding and activity, which might be impacted by making many mutations at the same time.

A

EPO	Disulfide bonds	Glycosylation sites	Variations in CD loop	Variations at C-terminus	Number of modifications
Human wildtype	C7–C161 C29–C33	N24, N38, N83, S126			
rsEPO-1	C7–C161 **C29–C33/139**	N24, N38, N83, S126			20
rsEPO-2	C7–C161 C29–C33	N24, N38, N83, S126			21
rsEPO-3	C7–C161 C29–C33	N24, N38, N83, S126, **N30, N88**			23
rsEPO-4	C7–C161 **C29–C88**	N24, N38, N83, S126			24
rsEPO-5	C7–C161 **C29–C88**	N24, N38, N83, S126, **N30**			25
rsEPO-6	C7–C161 **C29–C88**	N24, N38, N83, S126	**A125TP**		25
rsEPO-7	C7–C161 **C29–C88**	N24, N38, N83, S126	**SP120PL, A125TP**		28
rsEPO-8	C7–C161 **C29–C88**	N24, N38, N83, S126, **N125**	**A111V, K116R, I119T, A125N, A127T**		37
rsEPO-9	C7–C161 **C29–C88**	N24, N38, N83, S126, **N53, N165**	**A114V, K116R, L130A**	**DR165NSTGQ**	37
rsEPO-10	C7–C161 **C29–C139**	N24, N38, N83, S126, **N30, N53, N88, N165**	**A114V, K116R, L130S**	**DR165NSTGQ**	45

Amino acid numbering is based on human wildtype EPO without adding the insertions.

B
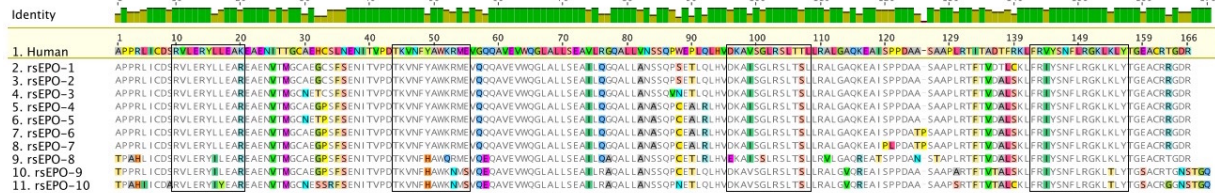
Figure 4.3: *(A) A list of rsEPO variants and modifications (bold) relative to human wild-type EPO. (B) Pairwise protein sequence alignment of rsEPO variants to human wild-type EPO. Highlighted amino acids indicate the differences between the variants and human EPO. Boxed regions are important for EPO–EPO-R interaction and activity. Sequences can be viewed in Appendix C.*

Lastly, additional *N*-linked glycans were introduced by mutating some amino acid sequences to the consensus sequence for *N*-linked glycosylation (N–X–S/T). Adding bulky sugar moieties interferes with the binding of both BCR and TCR to antigenic peptides by masking immune recognition epitopes in the protein moieties at both three-dimensional and one-dimensional levels. By affecting both T and B cell activation, this strategy holds two potential benefits. One is that if anti-drug antibodies form against the artificially glycosylated regions, they will not be cross-reactive with endogenous proteins. The other is that more long term and stronger immune responses, which require help from activated T cells, can be reduced. For example, the human immunodeficiency viruses (HIVs) decorate their envelope glycoprotein, gp120, with 24 *N*-linked oligosaccharides, such that antigen processing by proteases as well as antigen recognition by TCR

is sterically hindered by the attached sugar moieties [125, 132, 133, 134].

Based on these three factors that contribute to folding, stability, biological function, and immunogenicity, I designed 10 resurfaced EPO variants (rsEPO), as listed in Figure 4.3A. The first variant, rsEPO-1, is the simplest form: the consensus sequence of 24 mammalian EPOs, prioritizing non-human amino acid sequences (Figure 4.3). It contains 20 amino acid residues that differ from the original human sequence and an odd number of cysteines, as is the case in most mammals other than humans. In addition to the highly conserved Cys7–Cys161 bond, rsEPO-1 will form another disulfide bond between two out of the three cysteines that are spatially near one another (Cys29–Cys33, Cys29–Cys139, or Cys33–Cys139). Although this design allows rsEPO-1 to resemble EPO proteins from mammals other than humans, heterogeneous disulfide bond formation among three cysteine residues introduces batch-to-batch variability in drug formulation. The second variant, rsEPO-2, was designed to avoid such heterogeneity by mutating Cys139 back to the human sequence, arginine, so that only one additional disulfide bond would form in this region (Cys29–Cys33). Cys29–Cys33 was preferred to Cys29–Cys139 and Cys33–Cys139 because it resulted in a protein structure that was energetically more favorable. In protein structural models, the Cys29–Cys139 and Cys33–Cys139 bonds in mouse and pig EPO structures, respectively, hold together the AB loop and the beginning of helix D, creating a larger cavity that is not shielded from the solvent, compared to Cys29–Cys33 found in human EPO [124]. The fourth variant, rsEPO-4, was designed to include the optimized disulfide bond (Cys29–Cys88) instead of the human one as discussed above. Additional *N*-linked glycosylation sites were engineered into rsEPO-3, 5, 8, 9, and 10, and CD loop sequences were varied in rsEPO-6 through 10 (Figure 4.3A). In these variants, about 12.0–27.1% of the total EPO residues (20–45 out of 166 residues), including 12.3–29.5% of the solvent-exposed surface residues (16–39 out of 127–132 residues with solvent-exposed surface area of >6 Å^2), were mutated (Figure 4.3 and Figure 4.4).

4.3.2 Protein Structural Models of Resurfaced EPO Variants

Structural models of resurfaced EPO variants were generated using PyMOL and Swiss PDB Viewer. The X-ray crystal structure of the EPO–EPO-R complex (PDB ID: 1EER) was used as a starting template. This structure included some underlying mutations that were designed to
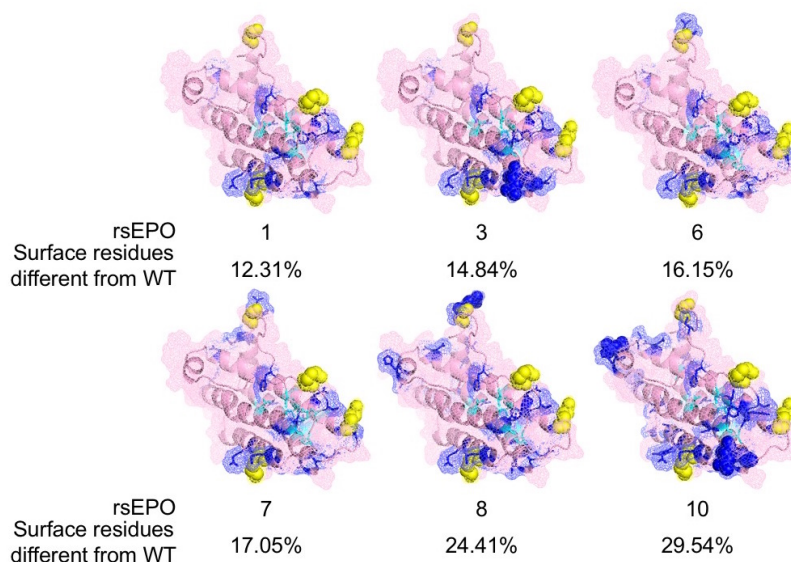
Figure 4.4: *Structural models of six representative rsEPO variants. Models were generated by sequence manipulation and energy minimization in PyMOL and Swiss PDB viewer, respectively, based on the structure of wild-type human EPO (PDB ID: 1EER). Modifications at the surface and internal residues are shown in blue and cyan, respectively. Glycosylation sites are shown as solid spheres. A fraction of surface residues that are different from wild-type EPO is calculated by dividing the number of mutated surface residues by the total number of surface residues. The surface residues are defined as residues with $>6Å^2$ of solvent-exposed area.*

help crystallization. All the *N*-linked glycosylation sites were mutated from asparagine to lysine (N24K, N38K, and N83K), and two proline residues at the beginning of the CD loop were mutated to asparagine and serine (P121N and P122S). When structural modeling was performed, the proline residues were put back in order to reflect the structure better but *N*-linked glycosylation sites were left as lysine residues because asparagine and lysine residues were considered similar enough to only minimally affect the entire structure. Then, all the mutations and insertions associated with each of the resurfaced EPO variants were introduced, and energy minimization was performed by Swiss PDB Viewer. In this process, the modified disulfide bonds were presented as separate, unbonded cysteine residues, and *N*- and *O*-linked oligosaccharides were not added to the glycosylation sites (Figure 4.4).

The structure of the overall protein backbone was not altered significantly with the resurfacing mutations and insertions, although incorporating modified disulfide bonds in these models would likely induce greater changes in the backbone structure. Most of the resurfacing modifications were located towards the ends of the helices, with the majority being at the opposite side from the N- and C-termini, and in the flexible CD loop (Figure 4.4). The outer regions of

the alpha helices were not modified very much because they are important for receptor binding and activation and therefore highly conserved among mammals [95, 131]. Surface variations were gradually increased from rsEPO-1 to rsEPO-10, resulting in about 12.3–29.5% of the solvent-exposed surface residues being mutated away from the human wild-type sequence (Figure 4.4). Engineered *N*-linked glycosylation patterns are expected to have a bigger impact than shown in the structural models. In the current models, oligosaccharides are not shown, and the amino acid residues at which glycosylation should occur are marked by solid spheres (Figure 4.4). When the sugar moieties are added to the structure, they would stretch about 20 Å away from the amino acid backbone and cover a decent fraction of the surrounding protein surface, resulting in a different surface appearance [49, 135].

4.3.3   Characterization of Resurfaced EPO Variants

Resurfaced EPO variants were produced recombinantly in mammalian cell culture and isolated by His-tag purification. Protein expression and secretion by HEK293F cells were verified by Western blotting of rsEPO in the cell culture supernatant 4 days after transfection. All 10 variants were successfully produced by HEK293F cells despite several sequence and structural modifications. Different variants appeared to have different molecular weights because of differences in sequence lengths and glycosylation patterns from various insertions and mutations. The variants with additional *N*-linked glycosylation sites appeared more smeared, as expected due to increased heterogeneity in glycosylation (Figure 4.5A).

Surprisingly, variants with more modifications resulted in significantly elevated expression levels (Figure 4.5A). Modifications that might be responsible for improved expression were deduced as follows. As rsEPO-1, 2, and 4 are designed by reversion to the consensus but have different disulfide bonds (Figure 4.3A), they serve as points of comparison for other resurfaced variants with similar mutational backgrounds. RsEPO-1 and 2 were expressed at similar levels, indicating that having heterogenous disulfide bonds among three cysteine residues (Cys29, Cys33, and Cys139) did not affect protein expression, compared to having the human disulfide bond (Cys29–Cys33). Expression levels of rsEPO-4 and 5 were also comparable, suggesting that additional *N*-linked glycosylation at position 30 (Asn30) did not affect protein expression.
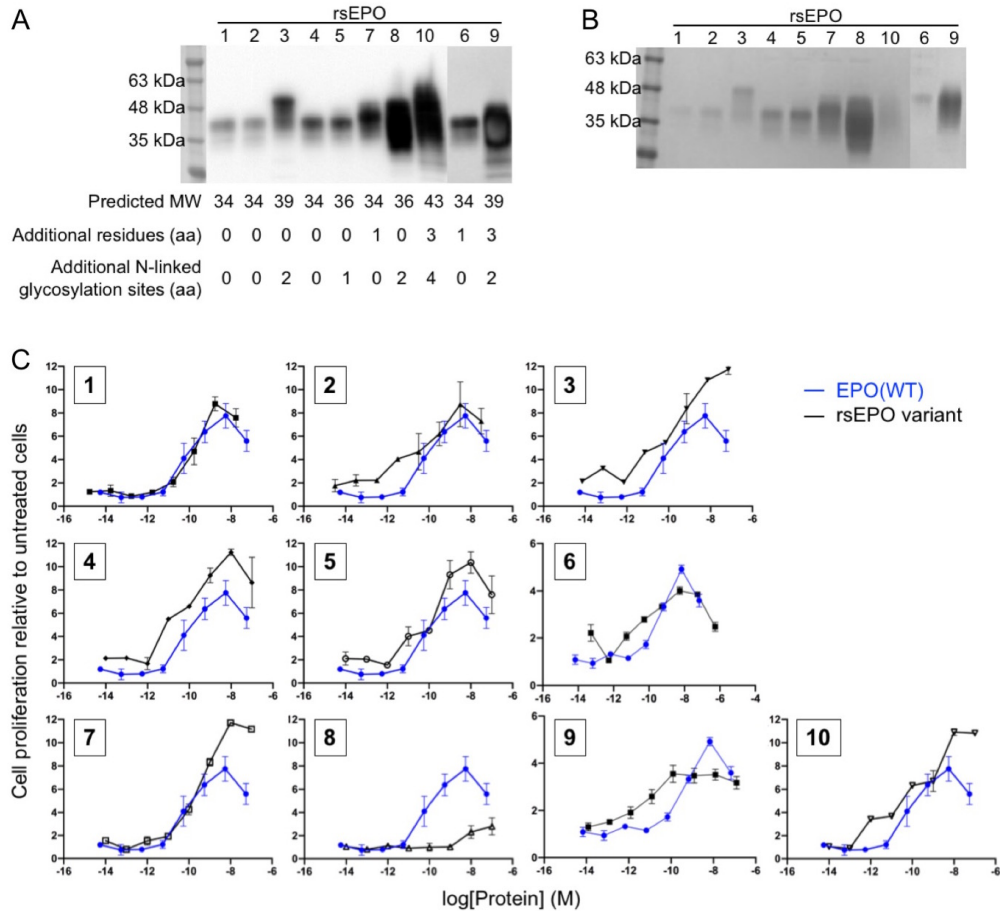
Figure 4.5: *Expression and purification of rsEPO variants. (A) Western blot of rsEPO proteins in HEK293F cell culture supernatant 4 days after transfection using anti-His$_6$-HRP antibody. Different molecular weights were observed due to different sequence lengths and glycosylation patterns. (B) Coomassie Blue staining of SDS–PAGE gels showing isolated proteins after His-tag purification. (C) In vitro erythropoietic activity of rsEPO variants were measured by cell proliferation assays. TF-1 cells were stimulated with increasing concentrations of EPO(WT) (Epogen for rsEPO-1–5, 7–8, 10; Aranesp for rsEPO-6, 9) or rsEPO variants for 72 hr and their growth was measured using MTS reagent. Data represent mean ± S.E.M. of three replicates.*

Interestingly, comparison between rsEPO-2 and 3 showed that having Asn88 in addition to Asn30 improved expression by 2.9-fold. Comparison of rsEPO-2 and 4 revealed that a certain disulfide bond may be better than others for folding. RsEPO-4 showed a 1.8-fold higher expression level than rsEPO-2, confirming Way *et al.*, 2005's claim that the Cys29–Cys88 disulfide bond is more favorable than the Cys29–Cys33 bond for protein folding and stability [124]. Modifications in the CD loop also improved protein expression. Both rsEPO-6 and 7, which contain the A125T mutation and an insertion of proline after this mutation, showed about a 1.6- to 1.9-fold increase in expression compared to rsEPO-4. The last three variants, rsEPO-8, 9, and 10, were expressed

at much higher levels than the other variants. These variants contain extensive modifications, including altered N-terminal sequences, the CD loop modifications, and extra glycosylation sites. One hypothesis for such high expression levels is that the alternative N-terminal sequences in these variants, which start with threonine instead of alanine, may facilitate the initiation of translation and initial folding of the nascent proteins. Another hypothesis is that mutations in the CD loop may stabilize otherwise very flexible and unstructured region. After His-tag purification, all variants were extracted with high purity (>99%) and acceptable yields (0.1–10 mg/L) (Figure 4.5B).

Proper folding and activity of these proteins were tested in a cell proliferation assay. TF-1 cells were treated with resurfaced variants or commercially available recombinant human EPO (Epogen® or Aranesp®, Amgen) for 72 hr. The commercial EPO stimulated TF-1 proliferation with EC_{50} of about 0.1–0.5 nM, as expected (Figure 4.5C). Nine out of ten rsEPO variants stimulated cell proliferation in a similar manner to commercial EPO, suggesting that these proteins were properly folded and biologically functional (Figure 4.5C). RsEPO-8 showed a slight increase in viable cell counts only at high concentrations (10–100 nM), indicating that it may not have folded properly, or that some of its mutations were deleterious to its bioactivity (Figure 4.5C). Given that rsEPO-8 is the variant that is most heavily mutated in the CD loop, it is speculated that mutations in that region, including a new *N*-linked glycosylation site right next to the native *O*-linked glycosylation site, might interfere with proper folding or receptor binding.

The resurfacing modifications in the other nine variants did not compromise folding or bioactivity of EPO. These modifications were incorporated into the final form of Targeted EPO, IH4-5-EPO(R150A), studied in Chapter 2. The goal was to ensure that Targeted EPO does not induce the generation of antibodies that cross-react with endogenous EPO and potentially lead to PRCA in treated patients. In these modified Targeted EPO fusion proteins, EPO contained both resurfacing modifications and the activity-weakening R150A mutation. These proteins were termed "FrsEPO," and were expressed, purified, and assayed for folding and activity. All nine FrsEPO proteins were successfully produced by HEK293F cells and purified by His-tag purification with acceptable yields (2–7 mg/mL). Their molecular weights were shifted by about 12 kDa relative to rsEPO proteins as expected from the size of the targeting element (IH4), but the relative size differences between FrsEPO variants remained the same as those observed among
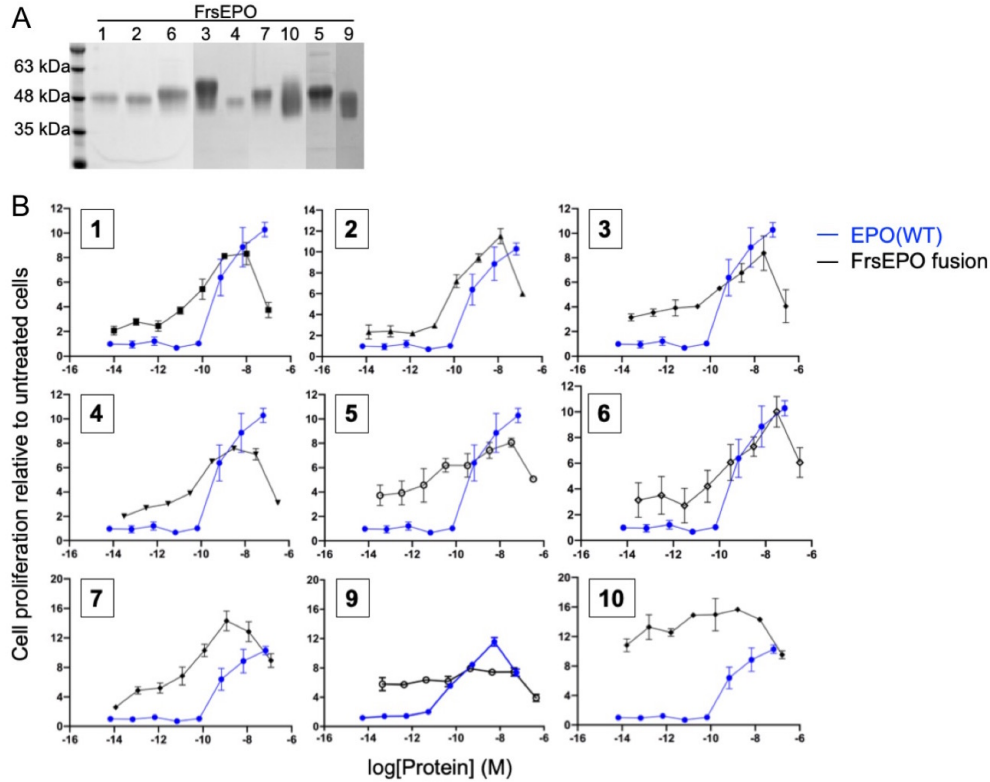
Figure 4.6: *Expression and purification of FrsEPO fusion proteins (IH4-5-rsEPO(R150A)). (A) Coomassie Blue staining of SDS–PAGE gels showing isolated proteins after His-tag purification. (B) Targeted erythropoietic activity of FrsEPO fusion proteins were measured by cell proliferation assays. TF-1 cells were stimulated with increasing concentrations of EPO(WT) (Epogen) or FrsEPO fusion proteins for 72 hr and their growth was measured using MTS reagent. Data represent mean ± S.E.M. of three replicates.*

rsEPO proteins (Figure 4.6A). Most of these fusion proteins stimulated proliferation of TF-1 cells with a similar or lower EC_{50} than the commercial EPO (Epogen®, Amgen). FrsEPO-5, 9 and 10 had high basal activity even at the lowest concentration tested (0.01 pM), suggesting that they may have been misfolded or aggregated (Figure 4.6B).

FrsEPO-1–4 and FrsEPO-6–7 will be further examined for their potential to generate antibodies that are cross-reactive with the endogenous EPO. The presence of shared antibody epitopes between the wild-type and resurfaced EPO will be predicted by monitoring the binding of these proteins to polyclonal antibodies, or anti-sera, for human or mouse wild-type EPO. The engineered proteins with least likelihood of cross-reactivity will be used in immunotoxicology studies in mice, in a similar manner to the previous experiment done with the original IH4-5-EPO(R150A) (Figure C.1).

4.4   Conclusions

Protein expression and immunogenicity are among the most common problems in developing protein therapeutics. Poor expression can raise the cost of production of a drug, or even deter a protein from being developed into a drug. Immunogenicity can affect both drug efficacy and safety. Although the consequences of immunogenicity vary depending on the protein itself, the disease being treated, and the condition of the patient, they can be as severe as the onset of life-threatening conditions and significant reduction of a patient's quality of life. Therefore, drug design should adopt engineering strategies to improve protein expression and reduce immunogenicity and its associated conditions.

EPO-induced PRCA, which depletes patients of their capability to produce new RBCs, is a good example of a case in which immune responses to recombinant protein drugs cause severe issues. In this chapter, I described an approach to mitigate the risk and thus the outcomes of immunogenicity and PRCA. More specifically, I analyzed protein sequence evolution among mammals and used protein structural information to rationally incorporate mutations. The mutated proteins are expected to present different surface epitopes, and therefore, prevent the generation of neutralizing antibodies that cross-react with the endogenous EPO. Despite having up to 27% of the total amino acid residues (45 out of 166) mutated, the altered proteins were expressed at significantly elevated levels, and maintained comparable levels of bioactivity to wild-type EPO. Structural models of resurfaced EPO variants revealed that while the structure of the protein backbone did not change significantly, about 12.3–29.5% of the surface residues would be modified and the addition of new glycosylation sites would augment the resurfacing effect by adding bulky sugar moieties. The present work demonstrates that mutations implemented in a rational manner can improve protein expression by stabilizing protein structure and folding as well as drug efficacy and safety by reducing their potential for generating neutralizing anti-drug antibodies without disrupting function.

4.5   Materials and Methods

Materials and Methods for cell culture of the 293-F cell line, DNA construction of EPO variants, protein expression and purification, in vitro measurement of erythropoietic activity by cell proliferation assays, and measurement of mouse reticulocytes and reticulated platelets by flow cytometry were described in Chapter 2.

Protein sequence alignment. Protein sequences of EPO from human and 23 other mammals were obtained from GenBank (see Appendix C for details). Resurfaced EPO variants were generated by manually changing amino acid residues from the human wild-type sequence. Pairwise alignments of EPO proteins were performed using Geneious (Biomatters, Auckland, New Zealand).

Structural Modeling. The X-ray crystal structure of human wild-type EPO (PDB ID: 1EER) was obtained from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB; rcsb.org). In 1EER, *N*-linked glycosylation sites of EPO were not reverted to asparagine and remained as lysine (N24K, N38K, and N83K). Two proline residues at the beginning of the CD loop were reverted to the original sequence, asparagine and serine (P121N and P122S). Structures of resurfaced EPO variants were generated by modifying amino acid sequences of human EPO in 1EER using PyMOL. Energy minimization was performed on the modified proteins using Swiss PDB Viewer.

Chapter 5

Conclusion

The field of protein engineering has expanded rapidly with growing needs for protein therapeutics over the last few decades. The first wave of protein drugs consisted of recombinantly synthesized cytokines, hormones and monoclonal antibodies that were essentially identical to proteins found in the human body. A new generation of protein drugs combines two or more protein parts to improve therapeutic effect, drug half-life, biodistribution, and target specificity. Chimeric activators, bispecific antibodies, and the engineered proteins that are part of chimeric antigen receptor T (CAR-T) cells are examples of this. Although the total sales revenue and the market share of protein therapeutics have increased dramatically, new drug approvals have not increased significantly. From 2001 to 2019, about a quarter of new drugs approved by the FDA each year were protein biologics [136]. The major challenge for new biologics lies in ensuring safety. Our limited understanding of complex biological systems may lead to unforeseen adverse effects associated with the intrinsic pleiotropy of the protein entity, on-target effects on non-target cells, or immune responses. Increasing knowledge in protein engineering will continue to drive smarter design strategies for safe and effective drugs, and this thesis work was performed in the hope of contributing to such efforts.

In this thesis, I have shown that protein-based synthetic biological systems can benefit from rational design strategies that are informed by the existing set of knowledge, such as protein sequence evolution, structure–function relationships, and molecular geometry of protein domains. Here, I used erythropoietin (EPO) as a model protein to demonstrate this principle, and explored three aspects of EPO that could be improved by applying rational protein engineering strategies.

77

In Chapter 2, tissue-targeted forms of EPO were constructed using a chimeric activator design and by analyzing the molecular geometry of the fusion proteins on the target cell surface. The resultant form of Targeted EPO, IH4-5-EPO(R150A), exhibited target cell-specific activity, reduced cell–cell crosslinking, and reduced systemic effects on bleeding time in mice. In Chapter 3, various flavors of EPO therapeutics with erythropoietic and/or tissue-protective functions were designed based on protein structure–function relationships and structural analyses of receptor–ligand complexes to serve various clinical needs. In Chapter 4, the surface residues of EPO were modified based on protein sequence evolution and protein structure–function relationships to prevent the formation of neutralizing antibodies that are cross-reactive with endogenous EPO. These three vignettes illustrate that the examination of protein structure and geometry enables us to manipulate protein functions and custom-design therapeutics with both a set of desired properties and reduced side effects.

These strategies and analytical thought processes are not just limited to the design of protein therapeutics. They are also very applicable to the broader scope of synthetic biology. Much of the field has focused on the engineering of gene expression systems to build and control artificial systems that perform desired functions, leveraging simple one-dimensional relationships between nucleic acid sequences and output proteins. In contrary, the nature has evolved to utilize three-dimensional structures, spatial geometry, and kinetics to coordinate various players using a limited amount of energy and resources in a complicated biological network. This certainly cannot be accomplished by relying on one-dimensional control of gene expression systems. The use of proteins in synthetic-biological systems shortens response time, and enables direct control over design outcomes by eliminating the need for transcription and translation. Furthermore, such synthetic systems can be refined more intricately by the manipulation of protein structures, combination of various functional domains from different proteins, and other protein engineering tricks.

One example of a protein-based synthetic-biological system is CAR-T cells, which induce killing of target cells, such as tumor cells, via synthetic receptors that bind to specific antigens on the target cell surface and signal for the release of cytotoxic agents. In such a system that requires cell-to-cell contact, the molecular geometry of the protein domains that compose the synthetic receptor becomes especially important. The length, orientation, or configuration of

the fusion protein can be modulated by similar approaches to those discussed in Chapter 2 to achieve favorable interactions. Another example is the rewiring of signaling pathways, wherein a hybrid of multiple protein domains from different signaling receptors can re-direct extracellular signals to different intracellular responses. Extraction of desired functions, modulation of signaling strength, and enhancement of specificity can be further refined by analyzing the structure–function relationships of each protein domain, as demonstrated in Chapter 3. The same three-dimensional, spatial, and quantitative thought processes may be applied to more complex protein circuit designs. In all of these systems, the potential for immune responses against synthetic proteins can be addressed by de-immunization techniques that utilize protein sequences and structures, as described in Chapter 4.

Rules have yet to be defined for the construction of synthetic-biological systems at a protein level. Despite our capabilities to recombinantly synthesize proteins, to decipher protein sequences, and to visualize protein structures, protein designs have primarily relied on empirical knowledge of "working" systems, with little effort to understand their rationale or to optimize them. For instance, antibody-based fusion proteins often employ the same configuration and a long linker, but the possibilities of that linker or configuration serving a different function than merely fusing two proteins were not important as long as the fusion protein "worked". The last three decades since the advent of the recombinant DNA technology have been dedicated to learning how to make protein therapeutics and manipulate them in a modular way. With the advances in the fields of synthetic biology and cell therapy, protein engineering has received more spotlight, and its application has expanded beyond the development of protein therapeutics. In order for protein engineering to help build more intricate and well-controlled systems, more attention should be given to how synthetic proteins may behave in a given space and what physicochemical properties of the protein should be modified to improve its performance. This thesis work shows that these considerations can facilitate the design process and help predict outcomes. I believe that with growing understanding of disease mechanisms and availability of three-dimensional protein structures, the progress of smarter protein design strategies is limited only by the creativity of synthetic biologists.

References

[1] K. Blankenship, "The top 20 drugs by 2018 U.S. sales," Jun 2019.

[2] M. J. Knauf, D. P. Bell, P. Hirtzer, Z.-P. Luo, J. D. Young, and N. V. Katre, "Relationship of effective molecular size to systemic clearance in rats of recombinant interleukin-2 chemically modified with water-soluble polymers," *Journal of Biological Chemistry*, vol. 263, no. 29, pp. 15064–15070, 1988.

[3] Z. Li, B.-F. Krippendorff, and D. K. Shah, "Influence of molecular size on the clearance of antibody fragments," *Pharmaceutical research*, vol. 34, no. 10, pp. 2131–2141, 2017.

[4] L. Thijs, C. Hack, R. Strack Van Schijndel, J. Nuijens, G. Wolbink, A. Eerenberg-Belmer, H. Van der Vall, and J. Wagstaff, "Activation of the complement system during immunotherapy with recombinant IL-2. Relation to the development of side effects.," *The Journal of Immunology*, vol. 144, no. 6, pp. 2419–2424, 1990.

[5] N. D. Vaziri and X. J. Zhou, "Potential mechanisms of adverse outcomes in trials of anemia correction with erythropoietin in chronic kidney disease," *Nephrology Dialysis Transplantation*, vol. 24, pp. 1082–1088, 2009.

[6] T. B. Drüeke, F. Locatelli, N. Clyne, K.-U. Eckardt, I. C. Macdougall, D. Tsakiris, H.-U. Burger, and A. Scherhag, "Normalization of hemoglobin level in patients with chronic kidney disease and anemia," *New England Journal of Medicine*, vol. 355, no. 20, pp. 2071–2084, 2006.

[7] A. K. Singh, L. Szczech, K. L. Tang, H. Barnhart, S. Sapp, M. Wolfson, and D. Reddan, "Correction of anemia with epoetin alfa in chronic kidney disease," *New England Journal of Medicine*, vol. 355, no. 20, pp. 2085–2098, 2006.

[8] M. A. Pfeffer, E. A. Burdmann, C.-Y. Chen, M. E. Cooper, D. De Zeeuw, K.-U. Eckardt, J. M. Feyzi, P. Ivanovich, R. Kewalramani, A. S. Levey, E. F. Lewis, J. B. McGill, J. J. McMurray, P. Parfrey, H.-H. Parving, G. Remuzzi, A. K. Singh, S. D. Solomon, and R. Toto, "A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease," *New England Journal of Medicine*, vol. 361, no. 21, pp. 2019–2032, 2009.

[9] M. Henke, R. Laszig, C. Rübe, U. Schäfer, K.-D. Haase, B. Schilcher, S. Mose, K. T. Beer, U. Burger, C. Dougherty, and H. Frommhold, "Erythropoietin to treat head and neck cancer patients with anaemia undergoing radiotherapy: randomised, double-blind, placebo-controlled trial," *The Lancet*, vol. 362, no. 9392, pp. 1255–1260, 2003.

[10] C. Hutmacher and D. Neri, "Antibody-cytokine fusion proteins: Biopharmaceuticals with immunomodulatory properties for cancer therapy," *Advanced drug delivery reviews*, vol. 141, pp. 67–91, 2019.

[11] N. D. Taylor, J. C. Way, P. A. Silver, and P. Cironi, "Anti-glycophorin single-chain Fv fusion to low-affinity mutant erythropoietin improves red blood cell-lineage specificity," *Protein Engineering, Design & Selection*, vol. 23, no. 4, pp. 251–260, 2010.

[12] H. F. Bunn, "Erythropoietin," *Cold Spring Harbor perspectives in medicine*, vol. 3, no. 3, p. a011619, 2013.

[13] J. C. Egrie, E. Dwyer, J. K. Browne, A. Hitz, and M. A. Lykos, "Darbepoetin alfa has a longer circulating half-life and greater *in vivo* potency than recombinant human erythropoietin," *Experimental hematology*, vol. 31, no. 4, pp. 290–299, 2003.

[14] R. S. Syed, S. W. Reid, C. Li, J. C. Cheetham, K. H. Aoki, B. Liu, H. Zhan, T. D. Osslund, A. J. Chirino, J. Zhang, J. Finer-Moore, S. Elliott, K. Sitney, B. A. Katz, D. J. Matthews, J. J. Wendoloski, J. Egrie, and R. M. Stroud, "Efficiency of signalling through cytokine receptors depends critically on receptor orientation," *Nature*, vol. 395, no. 6701, pp. 511–516, 1998.

[15] Y. L. Zhang, M. L. Radhakrishnan, X. Lu, A. W. Gross, B. Tidor, and H. F. Lodish, "Symmetric signaling by an asymmetric 1 erythropoietin: 2 erythropoietin receptor complex," *Molecular cell*, vol. 33, no. 2, pp. 266–274, 2009.

[16] J. S. Philo, K. H. Aoki, T. Arakawa, L. O. Narhi, and J. Wen, "Dimerization of the extracellular domain of the erythropoietin (epo) receptor by epo: one high-affinity and one low-affinity interaction," *Biochemistry*, vol. 35, no. 5, pp. 1681–1691, 1996.

[17] O. Livnah, E. A. Stura, S. A. Middleton, D. L. Johnson, L. K. Jolliffe, and I. A. Wilson, "Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation," *Science*, vol. 283, no. 5404, pp. 987–990, 1999.

[18] I. Remy, I. A. Wilson, and S. W. Michnick, "Erythropoietin receptor activation by a ligand-induced conformation change," *Science*, vol. 283, no. 5404, pp. 990–993, 1999.

[19] I. Moraga, G. Wernig, S. Wilmes, V. Gryshkova, C. P. Richter, W.-J. Hong, R. Sinha, F. Guo, H. Fabionar, T. S. Wehrman, P. Krutzik, S. Demharter, I. Plo, I. L. Weissman, P. Minary, R. Majeti, S. N. Constantinescu, J. Piehler, and K. C. Garcia, "Tuning cytokine receptor signaling by re-orienting dimer geometry with surrogate ligands," *Cell*, vol. 160, no. 6, pp. 1196–1208, 2015.

[20] T. D. Richmond, M. Chohan, and D. L. Barber, "Turning cells red: signal transduction mediated by erythropoietin," *Trends in cell biology*, vol. 15, no. 3, pp. 146–155, 2005.

[21] F.-K. Lin, S. Suggs, C.-H. Lin, J. K. Browne, R. Smalling, J. C. Egrie, K. K. Chen, G. M. Fox, F. Martin, and Z. Stabinsky, "Cloning and expression of the human erythropoietin gene," *Proceedings of the National Academy of Sciences*, vol. 82, no. 22, pp. 7580–7584, 1985.

[22] J. W. Eschbach, J. C. Egrie, M. R. Downing, J. K. Browne, and J. W. Adamson, "Correction of the anemia of end-stage renal disease with recombinant human erythropoietin," *New England Journal of Medicine*, vol. 316, no. 2, pp. 73–78, 1987.

[23] K. Kalantar-Zadeh, "History of erythropoiesis-stimulating agents, the development of biosimilars, and the future of anemia treatment in nephrology," *American journal of nephrology*, vol. 45, no. 3, pp. 235–247, 2017.

[24] V. Juneja, P. Keegan, J. E. Gootenberg, M. D. Rothmann, Y. L. Shen, K. Y. Lee, K. D. Weiss, and R. Pazdur, "Continuing reassessment of the risks of erythropoiesis-stimulating agents in patients with cancer," *Clinical cancer research*, vol. 14, no. 11, pp. 3242–3247, 2008.

[25] M.-E. Goebeler, S. Knop, A. Viardot, P. Kufer, M. S. Topp, H. Einsele, R. Noppeney, G. Hess, S. Kallert, A. Mackensen, *et al.*, "Bispecific T-cell engager (BiTE) antibody construct blinatumomab for the treatment of patients with relapsed/refractory non-Hodgkin lymphoma: final results from a phase I study," *Journal of Clinical Oncology*, vol. 34, no. 10, pp. 1104–1111, 2016.

[26] Z. L. Chang and Y. Y. Chen, "Cars: synthetic immunoreceptors for cancer therapy and beyond," *Trends in molecular medicine*, vol. 23, no. 5, pp. 430–450, 2017.

[27] P. M. Sondel and S. D. Gillies, "Current and potential uses of immunocytokines as cancer immunotherapy," *Antibodies*, vol. 1, no. 2, pp. 149–171, 2012.

[28] S. Kontos and J. A. Hubbell, "Improving protein pharmacokinetics by engineering erythrocyte affinity," *Molecular pharmaceutics*, vol. 7, no. 6, pp. 2141–2147, 2010.

[29] S. Kontos, I. C. Kourtis, K. Y. Dane, and J. A. Hubbell, "Engineering antigens for in situ erythrocyte binding induces T-cell deletion," *Proceedings of the National Academy of Sciences*, vol. 110, no. 1, pp. E60–E68, 2013.

[30] C. H. Villa, A. C. Anselmo, S. Mitragotri, and V. Muzykantov, "Red blood cells: Supercarriers for drugs, biologicals, and nanoparticles and inspiration for advanced delivery systems," *Advanced drug delivery reviews*, vol. 106, pp. 88–103, 2016.

[31] D. Spitzer, J. E. McDunn, S. Plambeck-Suess, P. S. Goedegebuure, R. S. Hotchkiss, and W. G. Hawkins, "A genetically encoded multifunctional TRAIL trimer facilitates cell-specific targeting and tumor cell killing," *Molecular cancer therapeutics*, vol. 9, no. 7, pp. 2142–2151, 2010.

[32] K. Tatzel, L. Kuroki, I. Dmitriev, E. Kashentseva, D. T. Curiel, S. P. Goedegebuure, M. A. Powell, D. G. Mutch, W. G. Hawkins, and D. Spitzer, "Membrane-proximal TRAIL species are incapable of inducing short circuit apoptosis signaling: Implications for drug development and basic cytokine biology," *Scientific reports*, vol. 6, no. 1, pp. 1–13, 2016.

[33] D. R. Burrill, A. Vernet, J. J. Collins, P. A. Silver, and J. C. Way, "Targeted erythropoietin selectively stimulates red blood cell expansion *in vivo*," *Proceedings of the National Academy of Sciences*, vol. 113, no. 19, pp. 5245–5250, 2016.

[34] A. Merry, E. Thomson, D. Anstee, and F. Stratton, "The quantification of erythrocyte antigen sites with monoclonal antibodies.," *Immunology*, vol. 51, no. 4, p. 793, 1984.

[35] A. Wickrema and J. D. Crispino, "Erythroid and megakaryocytic transformation," *Oncogene*, vol. 26, no. 47, pp. 6803–6815, 2007.

[36] A. Anagnostou, E. S. Lee, N. Kessimian, R. Levinson, and M. Steiner, "Erythropoietin has a mitogenic and positive chemotactic effect on endothelial cells.," *Proceedings of the National Academy of Sciences*, vol. 87, no. 15, pp. 5978–5982, 1990.

[37] L. Lifshitz, G. Tabak, M. Gassmann, M. Mittelman, and D. Neumann, "Macrophages as novel target cells for erythropoietin," *Haematologica*, vol. 95, no. 11, pp. 1823–1831, 2010.

[38] B. Gardner, S. Parsons, A. Merry, and D. Anstee, "Epitopes on sialoglycoprotein alpha: evidence for heterogeneity in the molecule.," *Immunology*, vol. 68, no. 2, p. 283, 1989.

[39] J. A. Chasis and N. Mohandas, "Red blood cell glycophorins," *Blood*, vol. 80, pp. 1869–1869, 1992.

[40] D. W. Knowles, J. A. Chasis, E. A. Evans, and N. Mohandas, "Cooperative action between band 3 and glycophorin A in human erythrocytes: immobilization of band 3 induced by antibodies to glycophorin a," *Biophysical Journal*, vol. 66, no. 5, pp. 1726–1732, 1994.

[41] M. Paulitschke, G. Nash, D. Anstee, M. Tanner, and W. Gratzer, "Perturbation of red blood cell membrane rigidity by extracellular ligands," *Blood*, vol. 86, no. 1, pp. 342–348, 1995.

[42] K. Giger, I. Habib, K. Ritchie, and P. S. Low, "Diffusion of glycophorin A in human erythrocytes," *Biochimica et Biophysica Acta (BBA)-Biomembranes*, vol. 1858, no. 11, pp. 2839–2845, 2016.

[43] I. Habib, D. Smolarek, C. Hattab, M. Grodecka, G. Hassanzadeh-Ghassabeh, S. Muyldermans, S. Sagan, C. Gutiérrez, S. Laperche, C. Le-Van-Kim, Y. C. Aronovicz, K. Wasniowska, S. Gangnard, and O. Bertrand, "VHH (nanobody) directed against human glycophorin A: A tool for autologous red cell agglutination assays," *Analytical biochemistry*, vol. 438, no. 1, pp. 82–89, 2013.

[44] J. Khoory, J. Estanislau, A. Elkhal, A. Lazaar, M. I. Melhorn, A. Brodsky, B. Illigens, I. Hamachi, Y. Kurishita, A. R. Ivanov, S. Shevkoplyas, N. I. Shapiro, and I. C. Ghiran, "Ligation of glycophorin a generates reactive oxygen species leading to decreased red blood cell function," *PLoS One*, vol. 11, no. 1, 2016.

[45] M. Piagnerelli, K. Z. BOUDJELTIA, B. Gulbis, M. Vanhaeverbeek, and J.-L. VINCENT, "Anemia in sepsis: the importance of red blood cell membrane changes," *Transfusion Alternatives in Transfusion Medicine*, vol. 9, no. 3, pp. 143–149, 2007.

[46] G. Paragh, S. M. Kumar, Z. Rakosy, S.-C. Choi, X. Xu, and G. Acs, "RNA interference-mediated inhibition of erythropoietin receptor expression suppresses tumor growth and invasiveness in A2780 human ovarian carcinoma cells," *The American journal of pathology*, vol. 174, no. 4, pp. 1504–1514, 2009.

[47] G. Acs, P. Acs, S. M. Beckwith, R. L. Pitts, E. Clements, K. Wong, and A. Verma, "Erythropoietin and erythropoietin receptor expression in human cancer," *Cancer research*, vol. 61, no. 9, pp. 3561–3565, 2001.

[48] S. K. Ballas, "Erythrocyte concentration and volume are inversely related," 1987.

[49] L. Kong, I. A. Wilson, and P. D. Kwong, "Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262," *Proteins: Structure, Function, and Bioinformatics*, vol. 83, no. 3, pp. 590–596, 2015.

[50] O. Schuster, G. Klich, V. Sinnwell, H. Kränz, H. Paulsen, and B. Meyer, "'Wave-type' structure of a synthetic hexaglycosylated decapeptide: A part of the extracellular domain of human glycophorin A," *Journal of biomolecular NMR*, vol. 14, no. 1, pp. 33–45, 1999.

[51] A. Kirkeby, L. Torup, L. Bochsen, M. Kjalke, K. Abel, K. Theilgaard-Monch, P. I. Johansson, S. E. Bjørn, J. Gerwien, and M. Leist, "High-dose erythropoietin alters platelet reactivity and bleeding time in rodents in contrast to the neuroprotective variant carbamyl-erythropoietin (CEPO)," *Thrombosis and Haemostasis*, vol. 99, no. 4, pp. 720–728, 2008.

[52] O. Bertrand, I. Habib, and D. Smolarek, "Fusion proteins and immunoconjugates and uses thereof which are specific for glycophorin A," Jan. 30 2018. US Patent 9,879,090.

[53] J. H. M. Cohen, W. Mahmoud, M. T. Libyh, N. Godin, A. Gimenez, T. Tabary, B. Donvito, D. Baty, and X. Dervillez, "Protein constructs designed for targeting and lysis of cells," Dec. 2 2014. US Patent 8,900,592.

[54] G. G. Lilley, P. J. Hudson, and C. J. Hillyard, "Reagent for agglutination assays," Dec. 09 1993. WO1993024630.

[55] I. Auffray, S. Marfatia, K. de Jong, G. Lee, C.-H. Huang, C. Paszty, M. J. Tanner, N. Mohandas, and J. A. Chasis, "Glycophorin A dimerization and band 3 interaction during erythroid membrane biogenesis: *in vivo* studies in human glycophorin A transgenic mice," *Blood, The Journal of the American Society of Hematology*, vol. 97, no. 9, pp. 2872–2878, 2001.

[56] M. I. Melhorn, A. S. Brodsky, J. Estanislau, J. A. Khoory, B. Illigens, I. Hamachi, Y. Kurishita, A. D. Fraser, A. Nicholson-Weller, E. Dolmatova, H. S. Duffy, and I. C. Ghiran, "CR1-mediated ATP release by human red blood cells promotes CR1 clustering and modulates the immune transfer process," *Journal of Biological Chemistry*, vol. 288, no. 43, pp. 31139–31153, 2013.

[57] B. Catimel, K. Wilson, and B. Kemp, "Kinetics of the autologous red cell agglutination test," *Journal of immunological methods*, vol. 165, no. 2, pp. 183–192, 1993.

[58] D. G. Haegert, "Demonstration of T4, T8, M1 and B7 determinants on human T cells with a rosette test: Implications for the assay specificity of monoclonal antibodies," *Journal of immunological methods*, vol. 82, no. 2, pp. 261–266, 1985.

[59] S. Elliott, A. Sinclair, H. Collins, L. Rice, and W. Jelkmann, "Progress in detecting cell-surface protein receptors: the erythropoietin receptor example," *Annals of hematology*, vol. 93, no. 2, pp. 181–192, 2014.

[60] P. Ghezzi and M. Brines, "Erythropoietin as an antiapoptotic, tissue-protective cytokine," *Cell Death & Differentiation*, vol. 11, no. 1, pp. S37–S44, 2004.

[61] S. Masuda, M. Okano, K. Yamagishi, M. Nagao, M. Ueda, and R. Sasaki, "A novel site of erythropoietin production. Oxygen-dependent production in cultured rat astrocytes.," *Journal of Biological Chemistry*, vol. 269, no. 30, pp. 19488–19493, 1994.

[62] O. O. Ogunshola and A. Y. Bogdanova, "Epo and non-hematopoietic cells: what do we know?," in *Tissue-Protective Cytokines*, pp. 13–41, Springer, 2013.

[63] C. C. Hernández, C. F. Burgos, A. H. Gajardo, T. Silva-Grecchi, J. Gavilan, J. R. Toledo, and J. Fuentealba, "Neuroprotective effects of erythropoietin on neurodegenerative and ischemic brain diseases: the role of erythropoietin receptor," *Neural regeneration research*, vol. 12, no. 9, p. 1381, 2017.

[64] M. Brines, G. Grasso, F. Fiordaliso, A. Sfacteria, P. Ghezzi, M. Fratelli, R. Latini, Q.-w. Xie, J. Smart, C.-j. Su-Rick, E. Pobre, D. Diaz, D. Gomez, C. Hand, T. Coleman, and A. Cerami, "Erythropoietin mediates tissue protection through an erythropoietin and common β-subunit heteroreceptor," *Proceedings of the National Academy of Sciences*, vol. 101, no. 41, pp. 14907–14912, 2004.

[65] L. Cherian, J. C. Goodman, and C. Robertson, "Neuroprotection with erythropoietin administration following controlled cortical impact injury in rats," *Journal of Pharmacology and Experimental Therapeutics*, vol. 322, no. 2, pp. 789–794, 2007.

[66] Z. Yu, L. Tang, L. Chen, J. Li, W. Wu, and C. Hu, "Erythropoietin reduces brain injury after intracerebral hemorrhagic stroke in rats," *Molecular medicine reports*, vol. 8, no. 5, pp. 1315–1322, 2013.

[67] S. Robinson, J. L. Winer, L. A. Chan, A. Y. Oppong, T. R. Yellowhair, J. R. Maxwell, N. Andrews, Y. Yang, L. O. Sillerud, W. P. Meehan III, R. Mannix, J. L. Brigman, and L. L. Jantzie, "Extended erythropoietin treatment prevents chronic executive functional and microstructural deficits following early severe traumatic brain injury in rats," *Frontiers in neurology*, vol. 9, p. 451, 2018.

[68] H. Ehrenreich, M. Hasselblatt, C. Dembowski, L. Cepek, P. Lewczuk, M. Stiefel, H.-H. Rustenbeck, N. Breiter, S. Jacob, F. Knerlich, M. Bohn, W. Poser, E. Rüther, M. Kochen, O. Gefeller, C. Gleiter, T. C. Wessel, M. De Ryck, L. Itri, H. Prange, A. Cerami, M. Brines, and A.-L. Sirén, "Erythropoietin therapy for acute stroke is both safe and beneficial," *Molecular medicine*, vol. 8, no. 8, pp. 495–505, 2002.

[69] S. Aloizos, E. Evodia, S. Gourgiotis, E.-C. Isaia, C. Seretis, and G. J. Baltopoulos, "Neuroprotective effects of erythropoietin in patients with severe closed brain injury," *Turkish Neurosurgery*, vol. 25, no. 4, pp. 552–558, 2015.

[70] D. Ostrowski and R. Heinrich, "Alternative erythropoietin receptors in the nervous system," *Journal of clinical medicine*, vol. 7, no. 2, p. 24, 2018.

[71] M. Leist, P. Ghezzi, G. Grasso, R. Bianchi, P. Villa, M. Fratelli, C. Savino, M. Bianchi, J. Nielsen, J. Gerwien, P. Kallunki, A. K. Larsen, L. Helboe, S. Christensen, L. O. Pedersen, M. Nielsen, L. Torup, T. Sager, A. Sfacteria, S. Erbayraktar, Z. Erbayraktar, N. Gokmen, O. Yilmaz, C. Cerami-Hand, Q.-w. Xie, T. Coleman, A. Cerami, and M. Brines, "Derivatives of erythropoietin that are tissue protective but not erythropoietic," *Science*, vol. 305, no. 5681, pp. 239–242, 2004.

[72] Y. Hanazono, K. Sasaki, H. Nitta, Y. Yazaki, and H. Hirai, "Erythropoietin induces tyrosine phosphorylation of the *beta* chain of the GM-CSF receptor," *Biochemical and biophysical research communications*, vol. 208, no. 3, pp. 1060–1066, 1995.

[73] J. M. Murphy, S. C. Ford, U. M. Wiedemann, P. D. Carr, D. L. Ollis, and I. G. Young, "A novel functional epitope formed by domains 1 and 4 of the human common *beta*-subunit is involved in receptor activation by granulocyte macrophage colony-stimulating factor and interleukin 5," *Journal of Biological Chemistry*, vol. 278, no. 12, pp. 10572–10577, 2003.

[74] S. E. Broughton, T. R. Hercus, T. L. Nero, M. Dottore, B. J. McClure, U. Dhagat, H. Taing, M. A. Gorman, J. King-Scott, A. F. Lopez, and M. W. Parker, "Conformational changes in

the GM-CSF receptor suggest a molecular mechanism for affinity conversion and receptor signaling," *Structure*, vol. 24, no. 8, pp. 1271–1281, 2016.

[75] M. E. Chamorro, S. D. Wenker, D. M. Vota, D. C. Vittori, and A. B. Nesse, "Signaling pathways of cell proliferation are involved in the differential effect of erythropoietin and its carbamylated derivative," *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, vol. 1833, no. 8, pp. 1960–1968, 2013.

[76] K. S. Shing, S. E. Broughton, T. L. Nero, K. Gillinder, M. D. Ilsley, H. Ramshaw, A. F. Lopez, M. D. Griffin, M. W. Parker, A. C. Perkins, and U. Dhagat, "EPO does not promote interaction between the erythropoietin and beta-common receptors," *Scientific reports*, vol. 8, no. 12457, pp. 1–16, 2018.

[77] S. K. Gaddam, J. Cruz, and C. Robertson, "Erythropoietin and cytoprotective cytokines in experimental traumatic brain injury," in *Tissue-Protective Cytokines*, pp. 141–162, Springer, 2013.

[78] H. Ehrenreich, B. Fischer, C. Norra, F. Schellenberger, N. Stender, M. Stiefel, A.-L. Sirén, W. Paulus, K.-A. Nave, R. Gold, and C. Bartels, "Exploring recombinant human erythropoietin in chronic progressive multiple sclerosis," *Brain*, vol. 130, no. 10, pp. 2577–2588, 2007.

[79] R. Nirula, R. Diaz-Arrastia, K. Brasel, J. Weigelt, and K. Waxman, "Safety and efficacy of erythropoietin in traumatic brain injury patients: a pilot randomized trial," *Critical care research and practice*, vol. 2010, 2010.

[80] H. Ehrenreich, K. Weissenborn, H. Prange, D. Schneider, C. Weimar, K. Wartenberg, P. D. Schellinger, M. Bohn, H. Becker, M. Wegrzyn, P. Jähnig, M. Herrmann, M. Knauth, M. Bähr, W. Heide, A. Wagner, S. Schwab, H. Reichmann, G. Schwendemann, R. Dengler, A. Kastrup, and C. Bartels, "Recombinant human erythropoietin in the treatment of acute ischemic stroke," *Stroke*, vol. 40, no. 12, pp. e647–e656, 2009.

[81] M. Brines, N. S. Patel, P. Villa, C. Brines, T. Mennini, M. De Paola, Z. Erbayraktar, S. Erbayraktar, B. Sepodes, C. Thiemermann, P. Ghezzi, M. Yamin, C. C. Hand, Q.-w. Xie, T. Coleman, and A. Cerami, "Nonerythropoietic, tissue-protective peptides derived from the tertiary structure of erythropoietin," *Proceedings of the National Academy of Sciences*, vol. 105, no. 31, pp. 10925–10930, 2008.

[82] W. Campana, R. Misasi, and J. O'brien, "Identification of a neurotrophic sequence in erythropoietin," *International journal of molecular medicine*, vol. 1, no. 1, pp. 235–276, 1998.

[83] Z. Erbayraktar, S. Erbayraktar, O. Yilmaz, A. Cerami, T. Coleman, and M. Brines, "Nonerythropoietic tissue protective compounds are highly effective facilitators of wound healing," *Molecular medicine*, vol. 15, pp. 235–241, 2009.

[84] I. Ahmet, H.-J. Tae, M. Juhaszova, D. R. Riordon, K. R. Boheler, S. J. Sollott, M. Brines, A. Cerami, E. G. Lakatta, and T. M. I., "A small nonerythropoietic helix B surface peptide based upon erythropoietin structure is cardioprotective against ischemic myocardial damage," *Molecular medicine*, vol. 17, pp. 194–200, 2011.

[85] M. Swartjes, A. Morariu, M. Niesters, M. Brines, A. Cerami, L. Aarts, and A. Dahan, "ARA290, a peptide derived from the tertiary structure of erythropoietin, produces long-term relief of neuropathic pain: An experimental study in rats and *beta*-common receptor knockout mice," *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 115, no. 5, pp. 1084–1092, 2011.

[86] W. G. van Rijt, H. van Goor, R. J. Ploeg, and H. G. Leuvenink, "Erythropoietin-mediated protection in kidney transplantation: nonerythropoietic EPO derivatives improve function without increasing risk of cardiovascular events," *Transplant International*, vol. 27, no. 3, pp. 241–248, 2014.

[87] Y. Gan, J. Xing, Z. Jing, R. A. Stetler, F. Zhang, Y. Luo, X. Ji, Y. Gao, and G. Cao, "Mutant erythropoietin without erythropoietic activity is neuroprotective against ischemic brain injury," *Stroke*, vol. 43, no. 11, pp. 3071–3077, 2012.

[88] M. Collino, C. Thiemermann, A. Cerami, and M. Brines, "Flipping the molecular switch for innate protection and repair of tissues: Long-lasting effects of a non-erythropoietic small peptide engineered from erythropoietin," *Pharmacology & therapeutics*, vol. 151, pp. 32–40, 2015.

[89] S. Masuda, M. Nagao, K. Takahata, Y. Konishi, F. Gallyas, T. Tabira, and R. Sasaki, "Functional erythropoietin receptor of the cells with neural characteristics. Comparison with receptor properties of erythroid cells.," *Journal of Biological Chemistry*, vol. 268, no. 15, pp. 11208–11216, 1993.

[90] W. M. Pardridge, "Transport of small molecules through the blood-brain barrier: biology and methodology," *Advanced drug delivery reviews*, vol. 15, no. 1-3, pp. 5–36, 1995.

[91] B. Schuler, J. Vogel, B. Grenacher, R. A. Jacobs, M. Arras, and M. Gassmann, "Acute and chronic elevation of erythropoietin in the brain improves exercise performance in mice without inducing erythropoiesis," *The FASEB Journal*, vol. 26, no. 9, pp. 3884–3890, 2012.

[92] M. L. Brines, P. Ghezzi, S. Keenan, D. Agnello, N. C. De Lanerolle, C. Cerami, L. M. Itri, and A. Cerami, "Erythropoietin crosses the blood–brain barrier to protect against experimental brain injury," *Proceedings of the National Academy of Sciences*, vol. 97, no. 19, pp. 10526–10531, 2000.

[93] W. A. Banks, N. L. Jumbe, C. L. Farrell, M. L. Niehoff, and A. C. Heatherington, "Passage of erythropoietic agents across the blood–brain barrier: a comparison of human and murine erythropoietin and the analog darbepoetin alfa," *European journal of pharmacology*, vol. 505, no. 1–3, pp. 93–101, 2004.

[94] J. Lee, A. Vernet, K. Redfield, S. Lu, I. C. Ghiran, J. C. Way, and P. A. Silver, "Rational design of a bifunctional AND-gate ligand to modulate cell–cell interactions," *ACS Synthetic Biology*, vol. 9, no. 2, pp. 191–197, 2020.

[95] S. Elliott, T. Lorenzini, D. Chang, J. Barzilay, and E. Delorme, "Mapping of the active site of recombinant human erythropoietin," *Blood, The Journal of the American Society of Hematology*, vol. 89, no. 2, pp. 493–502, 1997.

[96] T. Kitamura, T. Tange, T. Terasawa, S. Chiba, T. Kuwaki, K. Miyagawa, Y. Piao, K. Miyazono, A. Urabe, and F. Takaku, "Establishment and characterization of a unique human cell

line that proliferates dependently on GM-CSF, IL-3, or erythropoietin," *Journal of cellular physiology*, vol. 140, no. 2, pp. 323–334, 1989.

[97] L. M. Marshall, A. Thureson-Klein, and R. C. Hunt, "Exclusion of erythrocyte-specific membrane proteins from clathrin-coated pits during differentiation of human erythroleukemic cells.," *Journal of Cell Biology*, vol. 98, no. 6, pp. 2055–2063, 1984.

[98] N. T. Ktistakis, D. Thomas, and M. G. Roth, "Characteristics of the tyrosine recognition signal for internalization of transmembrane surface glycoproteins.," *The Journal of cell biology*, vol. 111, no. 4, pp. 1393–1407, 1990.

[99] G. Fuh, B. C. Cunningham, R. Fukunaga, S. Nagata, D. V. Goeddel, and J. A. Wells, "Rational design of potent antagonists to the human growth hormone receptor," *Science*, vol. 256, no. 5064, pp. 1677–1680, 1992.

[100] M. Atanasova and A. Whitty, "Understanding cytokine and growth factor receptor activation mechanisms," *Critical reviews in biochemistry and molecular biology*, vol. 47, no. 6, pp. 502–530, 2012.

[101] S. M. Kallenberger, J. Beaudouin, J. Claus, C. Fischer, P. K. Sorger, S. Legewie, and R. Eils, "Intra- and interdimeric caspase-8 self-cleavage controls strength and timing of CD95-induced apoptosis," *Science Signaling*, vol. 7, no. 316, pp. ra23–ra23, 2014.

[102] F. Stomski, Q. Sun, C. Bagley, J. Woodcock, G. Goodall, R. Andrews, M. Berndt, and A. Lopez, "Human interleukin-3 (IL-3) induces disulfide-linked IL-3 receptor alpha-and beta-chain heterodimerization, which is required for receptor activation but not high-affinity binding," *Molecular and cellular biology*, vol. 16, no. 6, pp. 3035–3046, 1996.

[103] P. T. Jubinsky, O. I. Krijanovski, D. G. Nathan, J. Tavernier, and C. A. Sieff, "The *beta* chain of the interleukin-3 receptor functionally associates with the erythropoietin receptor," *Blood, The Journal of the American Society of Hematology*, vol. 90, no. 5, pp. 1867–1873, 1997.

[104] S. Kusano, M. Kukimoto-Niino, N. Hino, N. Ohsawa, M. Ikutani, S. Takaki, K. Sakamoto, M. Hara-Yokoyama, M. Shirouzu, K. Takatsu, *et al.*, "Structural basis of interleukin-5 dimer recognition by its *α* receptor," *Protein Science*, vol. 21, no. 6, pp. 850–864, 2012.

[105] A. F. Lopez, M. J. Elliott, J. Woodcock, and M. A. Vadas, "GM-CSF, IL-3 and IL-5: cross-competition on human haemopoietic cells," *Immunology today*, vol. 13, no. 12, pp. 495–500, 1992.

[106] D. A. Rozwarski, K. Diederichs, R. Hecht, T. Boone, and P. A. Karplus, "Refined crystal structure and mutagenesis of human granulocyte-macrophage colony-stimulating factor," *Proteins: Structure, Function, and Bioinformatics*, vol. 26, no. 3, pp. 304–313, 1996.

[107] J. C. Egrie and J. K. Browne, "Development and characterization of novel erythropoiesis stimulating protein (nesp)," *British journal of cancer*, vol. 84, no. 1, pp. 3–10, 2001.

[108] A. Nichol, C. French, L. Little, S. Haddad, J. Presneill, Y. Arabi, M. Bailey, D. J. Cooper, J. Duranteau, O. Huet, *et al.*, "Erythropoietin in traumatic brain injury (epo-tbi): a double-blind randomised controlled trial," *The Lancet*, vol. 386, no. 10012, pp. 2499–2506, 2015.

[109] R. J. Boado, E. K.-W. Hui, J. Z. Lu, and W. M. Pardridge, "Drug targeting of erythropoietin across the primate blood-brain barrier with an igg molecular trojan horse," *Journal of Pharmacology and Experimental Therapeutics*, vol. 333, no. 3, pp. 961–969, 2010.

[110] R. Kunert and D. Reinhart, "Advances in recombinant antibody manufacturing," *Applied microbiology and biotechnology*, vol. 100, no. 8, pp. 3451–3461, 2016.

[111] V. Jawa, L. P. Cousens, M. Awwad, E. Wakshull, H. Kropshofer, and A. S. De Groot, "T-cell dependent immunogenicity of protein therapeutics: preclinical assessment and mitigation," *Clinical immunology*, vol. 149, no. 3, pp. 534–555, 2013.

[112] A. S. De Groot and D. W. Scott, "Immunogenicity of protein therapeutics," *Trends in immunology*, vol. 28, no. 11, pp. 482–490, 2007.

[113] I. M. Roitt, *Essential immunology.* Blackwell, 8 ed., 1994.

[114] J. Rossert, N. Casadevall, and K.-U. Eckardt, "Anti-erythropoietin antibodies and pure red cell aplasia," *Journal of the American Society of Nephrology*, vol. 15, no. 2, pp. 398–406, 2004.

[115] I. C. Macdougall, S. D. Roger, A. De Francisco, D. J. Goldsmith, H. Schellekens, H. Ebbers, W. Jelkmann, G. London, N. Casadevall, W. H. Hörl, D. M. Kemeny, and C. Pollock, "Antibody-mediated pure red cell aplasia in chronic kidney disease patients receiving erythropoiesis-stimulating agents: new insights," *Kidney international*, vol. 81, no. 8, pp. 727–732, 2012.

[116] C. L. Bennett, S. Luminari, A. R. Nissenson, M. S. Tallman, S. A. Klinge, N. McWilliams, J. M. McKoy, B. Kim, E. A. Lyons, S. M. Trifilio, D. W. Raisch, A. M. Evens, T. M. Kuzel, G. T. Schumock, S. M. Belknap, F. Locatelli, J. Rossert, and N. Casadevall, "Pure red-cell aplasia and epoetin therapy," *New England Journal of Medicine*, vol. 351, no. 14, pp. 1403–1408, 2004.

[117] N. Casadevall, J. Nataf, B. Viron, A. Kolta, J.-J. Kiladjian, P. Martin-Dupont, P. Michaud, T. Papo, V. Ugo, I. Teyssandier, B. Varet, and P. Mayeux, "Pure red-cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin," *New England Journal of Medicine*, vol. 346, no. 7, pp. 469–475, 2002.

[118] C. L. Bennett, D. Cournoyer, K. R. Carson, J. Rossert, S. Luminari, A. M. Evens, F. Locatelli, S. M. Belknap, J. M. McKoy, E. A. Lyons, B. Kim, R. Sharma, S. Costello, E. B. Toffelmire, G. A. Wells, H. A. Messner, P. R. Yarnold, S. M. Trifilio, D. W. Raisch, T. M. Kuzel, A. Nissenson, L.-C. Lim, M. S. Tallman, and N. Casadevall, "Long-term outcome of individuals with pure red cell aplasia and antierythropoietin antibodies in patients treated with recombinant epoetin: a follow-up report from the Research on Adverse Drug Events and Reports (RADAR) project," *Blood, The Journal of the American Society of Hematology*, vol. 106, no. 10, pp. 3343–3347, 2005.

[119] J. M. McKoy, R. E. Stonecash, D. Cournoyer, J. Rossert, A. R. Nissenson, D. W. Raisch, N. Casadevall, and C. L. Bennett, "Epoetin-associated pure red cell aplasia: past, present, and future considerations," *Transfusion*, vol. 48, no. 8, pp. 1754–1762, 2008.

[120] H. Schellekens, "Immunogenicity of therapeutic proteins," *Nephrology Dialysis Transplantation*, vol. 18, no. 7, pp. 1257–1259, 2003.

[121] A. J. Chirino, M. L. Ary, and S. A. Marshall, "Minimizing the immunogenicity of protein therapeutics," *Drug discovery today*, vol. 9, no. 2, pp. 82–90, 2004.

[122] S. Tangri, B. R. Mothé, J. Eisenbraun, J. Sidney, S. Southwood, K. Briggs, J. Zinckgraf, P. Bilsel, M. Newman, R. Chesnut, C. LiCalsi, and A. Sette, "Rationally engineered therapeutic proteins with reduced immunogenicity," *The Journal of Immunology*, vol. 174, no. 6, pp. 3187–3196, 2005.

[123] S. Delluc, G. Ravot, and B. Maillere, "Quantification of the preexisting CD4 T-cell repertoire specific for human erythropoietin reveals its immunogenicity potential," *Blood, The Journal of the American Society of Hematology*, vol. 116, no. 22, pp. 4542–4545, 2010.

[124] J. C. Way, S. Lauder, B. Brunkhorst, S.-M. Kong, A. Qi, G. Webster, I. Campbell, S. McKenzie, Y. Lan, B. Marelli, L. A. Nguyen, S. Degon, K.-M. Lo, and S. D. Gillies, "Improvement of Fc–erythropoietin structure and pharmacokinetics by modification at a disulfide bond," *Protein Engineering Design and Selection*, vol. 18, no. 3, pp. 111–118, 2005.

[125] J. N. Reitter, R. E. Means, and R. C. Desrosiers, "A role for carbohydrates in immune evasion in AIDS," *Nature medicine*, vol. 4, no. 6, pp. 679–684, 1998.

[126] J. L. Ottesen, P. Nilsson, J. Jami, D. Weilguny, M. Dührkop, D. Bucchini, S. Havelund, and J. Fogh, "The potential immunogenicity of human insulin and insulin analogues evaluated in a transgenic mouse model," *Diabetologia*, vol. 37, no. 12, pp. 1178–1185, 1994.

[127] K. W. Woodburn, Q. Fan, S. Winslow, M.-j. Chen, R. B. Mortensen, N. Casadevall, R. B. Stead, and P. J. Schatz, "Hematide is immunologically distinct from erythropoietin and corrects anemia induced by antierythropoietin antibodies in a rat pure red cell aplasia model," *Experimental hematology*, vol. 35, no. 8, pp. 1201–1208, 2007.

[128] I. C. Macdougall, J. Rossert, N. Casadevall, R. B. Stead, A.-M. Duliege, M. Froissart, and K.-U. Eckardt, "A peptide-based erythropoietin-receptor agonist for pure red-cell aplasia," *New England Journal of Medicine*, vol. 361, no. 19, pp. 1848–1855, 2009.

[129] I. C. Macdougall, R. Provenzano, A. Sharma, B. S. Spinowitz, R. J. Schmidt, P. E. Pergola, R. I. Zabaneh, S. Tong-Starksen, M. R. Mayo, H. Tang, K. R. Polu, A.-M. Duliege, and S. Fishbane, "Peginesatide for anemia in patients with chronic kidney disease not receiving dialysis," *New England Journal of Medicine*, vol. 368, no. 4, pp. 320–332, 2013.

[130] M. Christie, D. Peritt, R. M. Torres, T. W. Randolph, and J. F. Carpenter, "The role of protein excipient in driving antibody responses to erythropoietin," *Journal of pharmaceutical sciences*, vol. 104, no. 12, pp. 4041–4055, 2015.

[131] D. Wen, J.-P. Boissel, M. Showers, B. C. Ruch, and H. F. Bunn, "Erythropoietin structure-function relationships. identification of functionally important domains.," *Journal of Biological Chemistry*, vol. 269, no. 36, pp. 22839–22846, 1994.

[132] M. A. Wolfert and G.-J. Boons, "Adaptive immune activation: glycosylation does matter," *Nature chemical biology*, vol. 9, no. 12, p. 776, 2013.

[133] A. M. Chapman and B. R. McNaughton, "Scratching the surface: resurfacing proteins to endow new properties and function," *Cell chemical biology*, vol. 23, no. 5, pp. 543–553, 2016.

[134] D. J. Vigerust and V. L. Shepherd, "Virus glycosylation: role in virulence and immune interactions," *Trends in microbiology*, vol. 15, no. 5, pp. 211–218, 2007.

[135] V. Slynko, M. Schubert, S. Numao, M. Kowarik, M. Aebi, and F. H.-T. Allain, "NMR structure determination of a segmentally labeled glycoprotein using in vitro glycosylation," *Journal of the American Chemical Society*, vol. 131, no. 3, pp. 1274–1281, 2009.

[136] A. Mullard, "2019 fda drug approvals," *Nature Reviews Drug Discovery*, vol. 19, pp. 79–84, 2020.

[137] R. Shogren, T. A. Gerken, and N. Jentoft, "Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: light-scattering studies of ovine submaxillary mucin," *Biochemistry*, vol. 28, no. 13, pp. 5525–5536, 1989.

[138] J. Pieper, K.-H. Ott, and B. Meyer, "Stabilization of the T1 fragment of glycophorin A(N) through interactions with N- and O-linked glycans," *Nature Structural Biology*, vol. 3, no. 3, pp. 228–232, 1996.

[139] A. Robinson-Mosher, T. Shinar, P. A. Silver, and J. Way, "Dynamics simulations for engineering macromolecular interactions," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 2, p. 025110, 2013.

[140] M. Reth, "Antigen receptors on B lymphocytes," *Annual review of immunology*, vol. 10, no. 1, pp. 97–121, 1992.

[141] J. A. Chasis, M. E. Reid, R. H. Jensen, and N. Mohandas, "Signal transduction by glycophorin A: role of extracellular and cytoplasmic domains in a modulatable process.," *The Journal of cell biology*, vol. 107, no. 4, pp. 1351–1357, 1988.

[142] W. L. Bigbee, M. Vanderlaan, S. S. Fong, and R. H. Jensen, "Monoclonal antibodies specific for the M- and N-forms of human glycophorin A," *Molecular immunology*, vol. 20, no. 12, pp. 1353–1362, 1983.

[143] E. T. Mack, P. W. Snyder, R. Perez-Castillejos, B. Bilgiçer, D. T. Moustakas, M. J. Butte, and G. M. Whitesides, "Dependence of avidity on linker length for a bivalent ligand–bivalent receptor model system," *Journal of the American Chemical Society*, vol. 134, no. 1, pp. 333–345, 2012.

Appendix A

Supporting Information for Chapter 2

A.1 Sequences of Anti-Glycophorin A Antibodies, EPO, and Linkers

1. DNA sequences

- IH4

```
CAGGTCCAAC TGCAGGAGAG CGGCGGGGGG TCAGTTCAGG CGGGGGGGAG TCTGCGGTTG
AGCTGCGTAG CTTCAGGCTA CACTGACAGC ACCTACTGCG TGGGATGGTT TCGGCAGGCA
CCCGGCAAGG AACGAGAGGG CGTTGCACGG ATCAACACTA TCTCCGGTCG GCCTTGGTAC
GCAGATAGTG TTAAGGGACG GTTTACTATT AGTCAGGATA ACTCTAAGAA TACCGTC TAC
CTTCAGATGA ATAGCCTGAA ACCGGAAGAC ACGGCTATTT ACTATTGCAC CCTTACAACT
GCCAACAGCA GAGGGTTTTG TTCTGGGGGA TATAACTACA AAGGACAGGG G ACC CAAGTC
ACTGTCAGC
```

- IH4*

```
CAGGTCCAAC TGCAAGAGAG CGGAGGAGGG TCTGTTCAAG CTGGCGGTTC CCTCCGGCTT
TCTTGCGTGG CGTCAGGCTA TACTGACAGC ACATACTGCG TGGGCTGGTT CAGGCAGGCC
CCTGGAAAGG AGCGCGAGGG CGTAGCCCGC ATAAATACTA TATCTGGCAG ACCGTGGTAC
GCTGACAGCG TGAAGGGACG GTTTACAATC AGTCAAGATA ACTCTAAAAA CACCGTG TTT
CTTCAAATGA ATTCTTTGAA ACCCGAAGAT ACTGCCATCT ATTATTGCAC ACTTACGACC
GCGAACTCAC GCGGTTTTTG TAGCGGAGGA TATAACTATA AAGGGCAAGG GCAGGTAACT
GTATCC
```

- 10F7 (GenBank Accession No. KX026660–3)

```
CAAGTTAAGT TGCAACAATC TGGTGCTGAA TTGGTTAAGC CAGGTGCTTC TGTTAAGTTG
TCTTGTAAGG CTTCTGGTTA CACCTTCAAC TCTTACTTTA TGCATTGGAT GAAGCAAAGA
CCAGTTCAAG GTTTGGAATG GATTGGTATG ATTAGACCAA ACGGTGGTAC TACCGATTAC
AACGAGAAGT TTAAGAACAA GGCTACTTTG ACTGTTGATA AGTCCTCTAA CACTGCTTAC
ATGCAATTGA ACTCTTTGAC TTCTGGTGAT TCTGCTGTTT ACTACTGTGC TAGATGGGAA
GGTTCTTACT ACGCTTTGGA TTACTGGGGT CAAGGTACCA CTGTTACTGT TTCTTCC
GGTGGAGGTG GATCTGGTGG TGGAGGATCT TCAGGAGGTG GTGGATCTTC C
GATATTGAGT TGACTCAATC TCCAGCTATT ATGTCTGCTA CCTTGGGTGA GAAGGTTACT
```

ATGACTTGTA GAGCTTCATC TAACGTTAAG TACATGTACT GGTACCAACA GAAGTCTGGT
GCTTCTCCAA AGTTGTGGAT TTACTACACT TCTAACTTGG CTTCTGGTGT TCCAGGTAGA
TTTTCTGGTT CAGGTTCTGG TACTTCCTAC TCTTTGACTA TTTCCTCTGT TGAAGCTGAA
GATGCTGCTA CTTACTACTG TCAACAATTC ACTTCTTCCC CATACACTTT TGGAGGAGGT
ACTAAGTTGG AAATCAAG

- 1C3

GAAGTCCGTC TGCTGGAAAG CGGGGGTGGT CCTGTGCAGC CTGGTGGGTC CCTGAAACTG
TCCTGTGCCG CAAGCGGGTT CGATTTTTCC AGATACTGGA TGAACTGGGT GAGGAGGGCT
CCAGGCAAGG GCCTGGAGTG GATCGGCGAG ATCAACCAGC AGTCCAGCAC CATCAATTAC
TCTCCCCCTC TGAAGGACAA GTTCATCATC AGCCGCGATA ACGCTAAGTC TACACTGTAT
CTGCAGATGA ATAAGGTGAG AAGCGAGGAC ACCGCCCTGT ACTATTGCGC TCGCCTGTCT
CTGACAGCCG CTGGCTTTGC CTATTGGGGC CAGGGCACCC TGGTGACAGT GTCTGCT
GGAGGAGGCT CTTCCGGAGG ATCCGGCAGC TCTGGCGGCT CCAGCTCTGG CGGC
GATATCGTGA TGAGCCAGTC TCCCTCCAGC CTGGCCGTGT CCGTGGGAGA GAAGGTGTCC
ATGAGCTGTA AGTCTTCCCA GTCTCTGTTC AACTCCAGAA CCCGCAAGAA TTACCTGACA
TGGTATCAGC AGAAGCCTGG CCAGAGCCCC AAGCCTCTGA TCTACTGGGC CAGCACCAGA
GAGTCTGGAG TGCCAGACCG CTTCACCGGC TCTGGATCCG GCACAGACTT CACCCTGACA
ATCAGCTCTG TGCAGGCCGA GGACCTGGCT GATTACTATT GCAAGCAGTC CTATAATCTG
AGGACCTTTG GCGGCGGCAC AAAGCTGGAG ATCAAG

- R18

CAGGTTAAAC TCCAGCAAAG TGGTGGCGGG CTCGTACAAC CAGGCGGTTC CCTCAAGTTG
TCCTGCGCCG CATCAGGGTT TACATTTAGC TCTTATGGTA TGTCTTGGTT TCGCCAGACG
CCTGACAAGC GACTCGAGCT GGTCGCTATC ATCAATAGTA ACGGAGGTAC TACATATTAT
CCCGACAGTG TGAAGGGGCG ATTTACCATT AGCCGGGACA ACGCCAAAAA TACACTGTAC
CTCCAGATGT CAAGCTTGAA ATCAGAAGAT ACGGCCATGT ACTATTGCGC TAGGGGGGGT
GGAAGGTGGC TTCTGGACTA TTATGGTCAG GGTACAACAG TGACAGTATC CTCC
GGTGGAGGTG GATCTGGTGG TGGAGGATCT TCAGGAGGTG GTGGATCTTC C
GACATAGAGC TTACACAATC TCCGTCATCA CTGGCAGTCT CAGCCGGGGA AAAAGTGACA
ATGTCATGCA AGTCAAGCCA GAGCGTTCTT TATTCATCTA ATCAGAAGAA CTACCTGGCA
TGGTATCAGC AGAAGCCGGG ACAGTCCCCT AAGCTCCTCA TCTACTGGGC AAGCACCAGG
GAATCCGGAG TGCCGGACAG GTTTACTGGG TCCGGTTCTG GGACGGATTT TACGCTTACG
ATATCAAGTG TCCAAGCTGA GGACCTCGCA GTATACTACT GTCACCAGTA CCTGTCTTCT
TCTACTTTTG GGGGTGGAAC GAAACTGGAA ATAAAA

- EPO(WT) (GenBank Accession No. KX026660–3)

GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTAAG
GAGGCTGAAA ATATTACTAC TGGTTGTGCT GAACATTGTT CTTTGAACGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTGG TCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTG TTTTGAGAGG TCAAGCTTTG
TTGGTTAATT CTTCTCAACC ATGGGAACCA TTGCAATTGC ATGTTGATAA GGCTGTTTCT
GGTTTGAGAT CTTTGACTAC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTATTA CTGCTGATAC TTTTAGAAAG
TTGTTTAGAG TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAACTG GTGATCGG

- EPO(R150A) (GenBank Accession No. KX026660–3)

GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTAAG

```
GAGGCTGAAA ATATTACTAC TGGTTGTGCT GAACATTGTT CTTTGAACGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTGG TCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTG TTTTGAGAGG TCAAGCTTTG
TTGGTTAATT CTTCTCAACC ATGGGAACCA TTGCAATTGC ATGTTGATAA GGCTGTTTCT
GGTTTGAGAT CTTTGACTAC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTATTA CTGCTGATAC TTTTAGAAAG
TTGTTTAGAG TTTACTCTAA CTTCTTG GCC  GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAACTG GTGATCGG
```

- EPO(K45D) (GenBank Accession No. KX026660–3)

```
GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTAAG
GAGGCTGAAA ATATTACTAC TGGTTGTGCT GAACATTGTT CTTTGAACGA GAATATTACT
GTTCCAGATA CT GAT GTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTGG TCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTG TTTTGAGAGG TCAAGCTTTG
TTGGTTAATT CTTCTCAACC ATGGGAACCA TTGCAATTGC ATGTTGATAA GGCTGTTTCT
GGTTTGAGAT CTTTGACTAC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTATTA CTGCTGATAC TTTTAGAAAG
TTGTTTAGAG TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAACTG GTGATCGG
```

- 5AA linker

```
TCTGGTGGTG GTTCC
```

- 7AA linker

```
GGAGGATCTG GTGGTGGTTC C
```

- 17AA linker

```
GGAGGATCCG GTGGTGGAGG ATCATCTGGT GGAGGATCTG GTGGTGGTTC C
```

- 29AA linker

```
GGAGGAAGTT CCGGTGGTGG ATCTTCTTCT GGAGGTGGAG GATCCGGTGG TGGAGGATCA
TCTGGTGGAG GATCTGGTGG TGGTTCC
```

- 35AA linker

```
GGTGGAGGTG GTTCCGGAGG AGGAAGTTCC GGTGGTGGAT CTTCTTCTGG AGGTGGAGGA
TCCGGTGGTG GAGGATCATC TGGTGGAGGA TCTGGTGGTG GTTCC
```

2. Protein sequences

- IH4 (US Patent 9879090)

```
QVQLQESGGG SVQAGGSLRL SCVASGYTDS TYCVGWFRQA PGKEREGVAR INTISGRPWY
ADSVKGRFTI SQDNSKNTV Y  LQMNSLKPED TAIYYCTLTT ANSRGFCSGG YNYKGQG T QV
TVS
```

- IH4* (US Patent 9879090)

```
QVQLQESGGG SVQAGGSLRL SCVASGYTDS TYCVGWFRQA PGKEREGVAR INTISGRPWY
ADSVKGRFTI SQDNSKNTV F  LQMNSLKPED TAIYYCTLTT ANSRGFCSGG YNYKGQGQVT
VS
```

94
```

- 10F7

  QVKLQQSGAE LVKPGASVKL SCKASGYTFN SYFMHWMKQR PVQGLEWIGM IRPNGGTTDY
  NEKFKNKATL TVDKSSNTAY MQLNSLTSGD SAVYYCARWE GSYYALDYWG QGTTVTVSS
  GGGGSGGGGS SGGGGSS
  DIELTQSPAI MSATLGEKVT MTCRASSNVK YMYWYQQKSG ASPKLWIYYT SNLASGVPGR
  FSGSGSGTSY SLTISSVEAE DAATYYCQQF TSSPYTFGGG TKLEIK

- 1C3 (Patent Application WO1993024630)

  EVRLLESGGG PVQPGGSLKL SCAASGFDFS RYWMNWVRRA PGKGLEWIGE INQQSSTINY
  SPPLKDKFII SRDNAKSTLY LQMNKVRSED TALYYCARLS LTAAGFAYWG QGTLVTVSA
  GGGSSGGSGS SGGSSSGG
  DIVMSQSPSS LAVSVGEKVS MSCKSSQSLF NSRTRKNYLT WYQQKPGQSP KPLIYWASTR
  ESGVPDRFTG SGSGTDFTLT ISSVQAEDLA DYYCKQSYNL RTFGGGTKLE IK

- R18 (US Patent 8900592)

  QVKLQQSGGG LVQPGGSLKL SCAASGFTFS SYGMSWFRQT PDKRLELVAI INSNGGTTYY
  PDSVKGRFTI SRDNAKNTLY LQMSSLKSED TAMYYCARGG GRWLLDYYGQ GTTVTVSS
  GGGGSGGGGS SGGGGSS
  DIELTQSPSS LAVSAGEKVT MSCKSSQSVL YSSNQKNYLA WYQQKPGQSP KLLIYWASTR
  ESGVPDRFTG SGSGTDFTLT ISSVQAEDLA VYYCHQYLSS STFGGGTKLE IK

- EPO(WT)

  APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA
  VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS
  PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDR

- EPO(R150A)

  APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA
  VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS
  PPDAASAAPL RTITADTFRK LFRVYSNFL<mark>A</mark> GKLKLYTGEA CRTGDR

- EPO(K45D)

  APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDT<mark>D</mark>VNFYA WKRMEVGQQA
  VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS
  PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDR

- 5AA linker

  SGGGS

- 7AA linker

  GGSGGGS

- 17AA linker

  GGSGGGGSSG GGSGGGS

- 29AA linker

  GGSSGGGSSS GGGGSGGGGS SGGGSGGGS

- 35AA linker

  GGGGSGGGSS GGGSSSGGGG SGGGGSSGGG SGGGS

## A.2 Supplementary Tables and Figures

**Table A.1:** *Summary of ability of diverse forms of Targeted EPO to stimulate proliferation of TF-1 cells in vitro. Bold indicates data shown in Figure 2.3B in Chapter 2.*

| Protein (V region–linker–EPO) | N | Activity *in vitro* Log($EC_{50}$) ± SD | $EC_{50}$ relative to epoetin alfa |
|---|---|---|---|
| **Epoetin alfa (Amgen)** | **16** | **-10.239 ± 0.345** | **1** |
| **Darbepoetin (Amgen)** | **8** | **-8.989 ± 0.135** | **14.67** |
| **EPO(WT)** | **3** | **-10.378 ± 0.528** | **0.87** |
| **EPO(R150A)** | **7** | **-8.078 ± 0.233** | **130.86** |
| **EPO(K45D)** | **3** | **-6.572 ± 0.754** | **7298** |
| **10F7-35-EPO(R150A)** | **4** | **-8.947 ± 0.206** | **16.71** |
| 10F7-29-EPO(R150A) | 6 | -9.320 ± 0.303 | 7.78 |
| 10F7-18-EPO(R150A) | 5 | -9.173 ± 0.297 | 10.89 |
| 10F7-17-EPO(R150A) | 6 | -9.163 ± 0.306 | 11.27 |
| 10F7-7-EPO(R150A) | 5 | -9.386 ± 0.253 | 6.39 |
| **10F7-5-EPO(R150A)** | **8** | **-9.378 ± 0.754** | **13.69** |
| **10F7-EPO(R150A)-HEL** | **4** | **-9.142 ± 0.501** | **16.02** |
| **10F7-EPO(R150A)-LEH** | **4** | **-8.648 ± 0.643** | **54.76** |
| **1C3-35-EPO(R150A)** | **6** | **-10.517 ± 0.471** | **0.81** |
| 1C3-EPO(R150A)-HEL | 4 | -10.682 ± 0.900 | 0.98 |
| 1C3-EPO(R150A)-LEH | 4 | -10.808 ± 0.527 | 0.30 |
| **R18-17-EPO(R150A)** | **2** | **-9.415 ± 0.409** | **6.44** |
| R18-7-EPO(R150A) | 2 | -10.074 ± 0.690 | 1.95 |
| **R18-5-EPO(R150A)** | **2** | **-9.495 ± 0.185** | **4.56** |
| R18-17-EPO(K45D) | 2 | -8.828 ± 0.786 | 39.24 |
| R18-7-EPO(K45D) | 1 | -8.494 ± N/A | 43.71 |
| R18-5-EPO(K45D) | 1 | -8.019 ± N/A | 130.58 |
| **IH4-35-EPO(R150A)** | **16** | **-10.385 ± 0.756** | **1.45** |
| IH4-17-EPO(R150A) | 5 | -10.243 ± 0.505 | 1.34 |
| IH4-7-EPO(R150A | 5 | -10.479 ± 0.645 | 1.13 |
| **IH4-5-EPO(R150A)** | **9** | **-10.410 ± 0.293** | **0.69** |
| **IH4-35-EPO(K45D)** | **5** | **-9.591 ± 0.364** | **4.70** |
| IH4-5-EPO(K45D) | 1 | -9.559 ± N/A | 3.76 |

A

Bleeding time (s): 600, 400, 200, 0

Experiment: 1 2 3 4 5 6 7 8 9 10 11 12 13

Legend:
- Vehicle
- Epoetin alfa
- Darbepoetin
- 10F7-35-EPO(R150A)
- IH4-35-EPO(R150A) 9 µg
- IH4-35-EPO(R150A) 1 µg
- IH4-5-EPO(R150A)

B

| Experiment | Human GPA | Dose groups |
|---|---|---|
| 1<br>2<br>3 | Non-transgenic | Vehicle<br>Epoetin alfa |
| 4<br>5 | Transgenic | Vehicle<br>Epoetin alfa<br>10F7-35-EPO(R150A) |
| 6<br>7<br>8 | Non-transgenic | Vehicle<br>10F7-35-EPO(R150A) |
| *9 | Transgenic | *Vehicle<br>Darbepoetin<br>IH4-35-EPO(R150A) 9 µg<br>IH4-35-EPO(R150A) 1 µg* |
| 10<br>11<br>12<br>13 | Transgenic | Vehicle<br>Darbepoetin<br>IH4-5-EPO(R150A) |

C

| Protein (V region–linker–EPO) | Dose (pmol) | Human GPA | N | Median bleeding time relative to vehicle ± SD | p-value | Significance |
|---|---|---|---|---|---|---|
| Epoetin alfa (Amgen) | 50 | Transgenic | 2 | 0.840 ± 0.190 | 0.0063 | *** |
| | 50 | Non-transgenic | 3 | 0.653 ± 0.074 | | |
| Darbepoetin (Amgen) | 50 | Transgenic | 4 | 0.828 ± 0.301 | 0.0207 | ** |
| 10F7-35-EPO(R150A) | 125 | Transgenic | 2 | 0.717 ± 0.098 | 0.0527 | * |
| | 125 | Non-transgenic | 3 | 1.257 ± 0.313 | 0.1750 | ns |
| IH4-5-EPO(R150A) | 40 | Transgenic | 4 | 1.028 ± 0.083 | 0.8015 | ns |
| *Darbepoetin (Amgen)* | *50* | *Transgenic* | *1* | *0.761 ± N/A* | *0.2855* | *ns* |
| *IH4-35-EPO(R150A)* | *180* | *Transgenic* | *1* | *0.533 ± N/A* | *0.1200* | *ns* |
| | *20* | *Transgenic* | *1* | *0.806 ± N/A* | *0.7815* | *ns* |

**Figure A.1:** *Summary of bleeding times after treating huGPA-transgenic or non-transgenic mice with various forms of non-targeted and targeted EPO. Mice received a single i.p. injection of darbepoetin or epoetin alfa (50 pmol; 1.8 µg), 10F7-EPO(R150A) (125 pmol; 9 µg), IH4-5-EPO(R150A) (40 pmol; 2 µg), or vehicle on day 0, and bleeding times were measured on day 1. In experiment 9 (italics, *), mice received two i.p. injections of darbepoetin (50 pmol; 1.8 µg), IH4-35-EPO(R150A) (180 pmol; 9 µg or 20 pmol; 1 µg), or vehicle on days 0 and 3, and bleeding times were measured on day 4. (A) Raw data from all 13 bleeding time experiments. Data represent median with interquartile range. (B) Description of each experiment shown in (A). (C) Bleeding times were normalized to the median of a vehicle control performed on the same day, to account for day-to-day variability. The normalized values were combined across experiments and Mann–Whitney test was performed. *p < 0.1; **p < 0.05; ***p < 0.01; ns: not significant.*

**Table A.2:** *Summary of epitopes and binding kinetics of GPA antibody fragments and EPO (wild-type and R150A mutant) studied in Chapter 2.*

| Protein | $k_{on}$ (M$^{-1}$s$^{-1}$) | $k_{off}$ (s$^{-1}$) | $K_D$ (nM) | Antibody form | GPA epitope | References |
|---|---|---|---|---|---|---|
| 10F7 | – | – | 95 | Fab | $_{34}$YAATP$_{38}$ | Chasis *et al*, 1988<br>Chasis & Mohandas, 1992<br>Catimel *et al.*, 1993 |
| 1C3 | – | – | 230 | Fab | – | Catimel *et al.*, 1993 |
| | – | – | 62 | scFv | – | Patent Application WO1994007921 |
| R18 | – | – | 25 | IgG | $_{49}$RTVY$_{52}$ | Gardner *et al.*, 1989 |
| | – | – | 400 | Fab | | |
| IH4 | $5.73 \times 10^5$ | 0.019 | 33.72 | Nanobody (VHH) | $_{52}$YPPE$_{55}$ | Habib *et al.*, 2013 |
| EPO(WT) | $3.9 \times 10^4$ | $2.1 \times 10^{-4}$ | 5.4 | N/A | N/A | Burrill *et al.*, 2016 |
| EPO(R150A) | $4.2 \times 10^4$ | $3.4 \times 10^{-3}$ | 81 | N/A | N/A | Burrill *et al.*, 2016 |



**Figure A.2:** *Ability of a form of Targeted EPO, R18-5-EPO(R150A), to promote RBC production in human GPA-transgenic mice. Compared to IH4-5-EPO(R150A), a higher dose of R18-5-EPO(R150A) is needed to induce comparable increase in reticulocytes.*

A

```
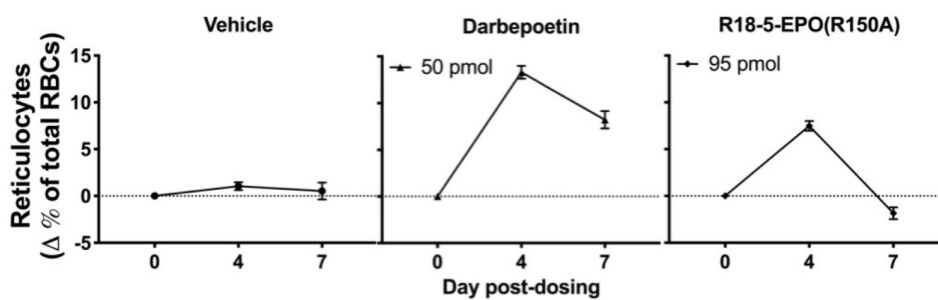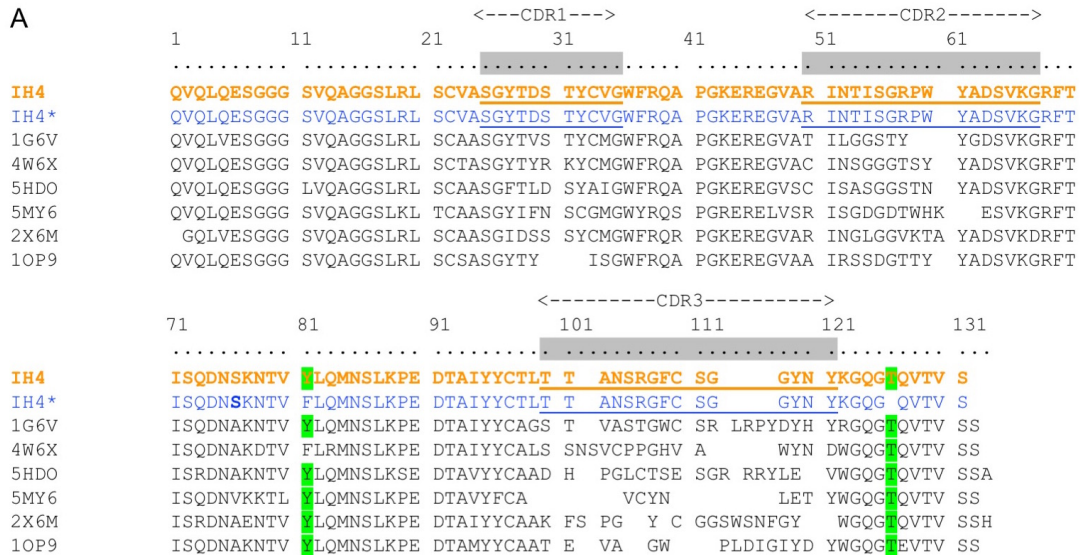                                   <---CDR1--->                    <-------CDR2------->
        1         11        21        31        41        51        61
        .......... .......... .......... .......... .......... .......... .......... ...

IH4     QVQLQESGGG SVQAGGSLRL SCVASGYTDS TYCVGWFRQA PGKEREGVAR INTISGRPW  YADSVKGRFT
IH4*    QVQLQESGGG SVQAGGSLRL SCVASGYTDS TYCVGWFRQA PGKEREGVAR INTISGRPW  YADSVKGRFT
1G6V    QVQLVESGGG SVQAGGSLRL SCAASGYTVS TYCMGWFRQA PGKEREGVAT ILGGSTY    YGDSVKGRFT
4W6X    QVQLQESGGG SVQAGGSLRL SCTASGYTYR KYCMGWFRQA PGKEREGVAC INSGGGTSY  YADSVKGRFT
5HDO    QVQLQESGGG LVQAGGSLRL SCAASGFTLD SYAIGWFRQA PGKEREGVSC ISASGGSTN  YADSVKGRFT
5MY6    QVQLQESGGG SVQAGGSLKL TCAASGYIFN SCGMGWYRQS PGRERELVSR ISGDGDTWHK  ESVKGRFT
2X6M     GQLVESGGG SVQAGGSLRL SCAASGIDSS SYCMGWFRQR PGKEREGVAR INGLGGVKTA YADSVKDRFT
1OP9    QVQLQESGGG SVQAGGSLRL SCSASGYTY    ISGWFRQA PGKEREGVAA IRSSDGTTY  YADSVKGRFT
```

```
                                            <---------CDR3---------->
        71        81        91        101       111       121       131
        .......... .......... .......... .......... .......... .......... .......... ...

IH4     ISQDNSKNTV YLQMNSLKPE DTAIYYCTLT T  ANSRGFC SG    GYN YKGQGTQVTV S
IH4*    ISQDNSKNTV FLQMNSLKPE DTAIYYCTLT T  ANSRGFC SG    GYN YKGQG QVTV S
1G6V    ISQDNAKNTV YLQMNSLKPE DTAIYYCAGS T  VASTGWC SR LRPYDYH YRGQGTQVTV SS
4W6X    ISQDNAKDTV FLRMNSLKPE DTAIYYCALS SNSVCPPGHV A      WYN DWGQGTQVTV SS
5HDO    ISRDNAKNTV YLQMNSLKSE DTAVYYCAAD H  PGLCTSE SGR RRYLE VWGQGTQVTV SSA
5MY6    ISQDNVKKTL YLQMNSLKPE DTAVYFCA      VCYN        LET YWGQGTQVTV SS
2X6M    ISRDNAENTV YLQMNSLKPE DTAIYYCAAK FS PG  Y C GGSWSNFGY  WGQGTQVTV SSH
1OP9    ISQDNAKNTV YLQMNSLKPE DTAMYYCAAT E  VA  GW     PLDIGIYD YWGQGTEVTV SS
```

B

1G6V: Dromedary nanobody to bovine carbonic anhydrase
4W6X: Llama nanobody to lectin domain of F18 fimbrial adhesin FedF
5HDO: Dromedary nanobody to urokinase-type plasminogen activator
5MY6: Dromedary nanobody to HER2
2X6M: Dromedary nanobody to alpha-synuclein
1OP9: Camelid nanobody to human lysozyme

Figure A.3: *Protein sequence alignments of IH4 and other nanobodies. (A) The original IH4 sequence (IH4*) was modified to match the consensus sequence. IH4* (blue) indicates the original protein sequence of the IH4 nanobody from the US Patent No. 9,879,090 [52]. Phe80 in the framework region 3 of IH4* is mutated to tyrosine (green highlight), and a threonine residue is inserted between Gly117 and Gln118 in the framework region 4 of IH4* (green highlight). The resulting sequence is shown as IH4 (orange, bold). Each dot indicates a position for a single amino acid residue. Numbers indicate amino acid positions every 10 residues. Complementarity-determining regions (CDRs) are shown with dots highlighted in gray and underlined sequences. Sequences between CDRs are framework regions. (B) PDB ID's and brief descriptions of six nanobodies used in sequence alignments. Six nanobodies with known sequences and structures were randomly chosen and compared.*

Figure A.4: *Phase-contrast images of RBC rosetting around EPO-R-bearing cells. Human RBCs treated with Targeted EPO variants were incubated with (B) A2780 or (B) MCF-7 cells. 200x magnification.*

Figure A.5: *Geometric model of the RBC surface. There are 800,000 GPA monomers densely packed on the RBC surface, whose area is about 140 µm² [34, 48]. The average distance between GPA molecules is about 130–190 Å. (A) When all GPA molecules are monomeric, the average distance between the two adjacent monomers is 132 Å. (B) When all GPA molecules dimerize, the average distance between the two adjacent dimers is 187 Å. O- and N-linked glycans extend 20 Å from the peptide backbone. About 90–100 Å between GPA molecules can be freely accessed by external molecules. (C) Summary of relevant numbers and equations to estimate the density and accessibility of GPA molecules on RBC.*

A.3   Estimation of Molecular Distances on Red Blood Cells

Molecular distances as schematically depicted in Figure 2.8 and as discussed in Chapter 2 were estimated as follows. The extracellular domain of GPA is generally described as an intrinsically unstructured protein, based on the fact that the first 45 amino acids of the protein include 17 *O*-glycosylation sites that are variably modified and one *N*-linked glycosylation site; these are so densely packed that formation of a hydrophobic core is likely precluded in this region, particularly for amino acids 1–33 [50, 137, 138].

The *O*-linked oligosaccharides are expected to extend at least about 20 Å from the peptide backbone. In the crystal and NMR structures of glycosylated proteins, the average distance between the peptide backbone and the farthest end of a single disaccharide unit is about 8–10 Å. Therefore, an *N*-/*O*-linked glycan is estimated to extend about 20–25 Å from the peptide backbone [49, 135].

Taken together, at least the first 33 amino acids are expected to form an extended coil, similar in conformation to a beta-strand with about 3–3.5 Å per amino acid. For example, in a three-dimensional GFP structure (PDB ID: 1C4F), the length of the Asn159–Gln172 beta strand is about 42 Å. Thus, the distance from the 10F7 epitope (roughly amino acid 33) to the N-terminus is likely at least 100 Å.

The length of a fully extended glycine–serine linker of 35 amino acids would also be about 100 Å. The flexible glycine–serine linker connecting the antibody and EPO elements is thought to form a random coil with a length distribution approximating that of a volume-excluded Brownian walk [139]. However, when the antibody element is bound, the EPO element likely explores all of the available space on a short timescale. Thus, if the EPO element binds to an EPO-R while the linker is in a completely extended conformation, that conformation will be stabilized. More precisely, when the antibody element is bound to GPA, the concentration of a single molecule of attached EPO in a sphere with a 100 Å radius is about 400 μM. Since the dissociation constant (K_D) of EPO(R150A) for EPO-R is about 81 nM [33], it is expected that if an EPO-R is found within the radius afforded by an extended linker, binding will occur.

The size, shape, and binding angle of the anti-GPA antibody V regions 10F7 and IH4 can be estimated from the structures of other scFv's and nanobodies, as well as from steric

considerations of how bound antibodies on the surface of B cells could interact with epitopes on GPA during antibody selection. It is important to recall that when a B cell expressing a particular antibody is being positively selected, it engages the antigen through membrane-bound antibodies that are brought in proximity [140]. The 10F7 and IH4 antibodies were both developed by immunization with whole human RBCs. The length of an antibody Fc region is about 80–90 Å, so that membrane-bound antibody could not be in the T conformation during antigen recognition. Instead, the Fab portion must be angled upward, away from the RBC membrane. As a result, the net displacement from the complementarity-determining regions (CDRs) of the V region(s) to the C-terminus of the V region(s) would need to be at least 10 Å and up to 45 Å (the length of a V region).

In the solved structures of EPO complexed with its receptor dimer (PDB ID: 1EER and 1CN4), EPO is roughly a prolate spheroid with its long axis parallel to the membrane of the EPO-R-bearing cell. In a configuration in which the EPO-Rs are maximally extended from the membrane, the N-terminus of EPO would be about 50 Å from the membrane. The N-terminus of EPO is at least about 25 Å from the most membrane-distal amino acids on the bound EPO-R, so if an EPO fusion protein is bound to GPA to a RBC, a linker would need to span this plus the distance from where it is anchored on the RBC. This reasoning assumes that EPO-R on one cell and GPA on another cell cannot interdigitate; this seems a reasonable assumption because of the glycosylation of GPA.

The IH4 nanobody element recognizes the sequence $_{52}$YPPE$_{55}$ in GPA [43]. This is about 20 amino acids C-terminal to the likely epitope of 10F7 [39, 141, 142].

Taken together, these estimates suggest that for 10F7-35-EPO(R150A), when 10F7 is bound to GPA on a RBC, the surface of EPO that binds to its receptor can protrude at least 10 Å beyond the N-terminus of GPA, even if this protein is in a maximally extended conformation. For IH4-5-EPO(R150A), the EPO element is about 150 Å closer to the RBC membrane, corresponding to a 20 amino acid more membrane proximal epitope, and a linker that is 30 amino acids shorter.

A.4   Estimation of Effective Concentrations on RBC Precursors

The choice of antibody element and linker length determines the effective concentration of EPO on the target cell surface, which consequently translates to potency and the minimal effective dose of the drug. The effective concentration of EPO is estimated by the amount of GPA-bound EPO within the volume of space in which EPO can move freely in the bound state. An antibody element with high affinity for GPA localizes the EPO element tightly to the cell surface, resulting in a high concentration of bound EPO at equilibrium. The length of the linker and the position of the binding epitope of the antibody element on GPA together affect the volume of space occupied by bound EPO. A short linker and a membrane-proximal antibody epitope limit the movement of GPA-bound EPO within a small region around the cell surface.

Table A.3: *Estimation of effective concentrations of IH4-5-EPO(R150A) and 10F7-35-EPO(R150A) on a RBC precursor cell.*

Equation	Description	IH4-5-EPO(R150A)	10F7-35-EPO(R150A)	References
(a)	[Targeted EPO]$_0$ in tissue extracellular space	1 nM (assumption)		
(b)	Avogadro's number	6.022×10^{23}		
(c)	GPA binding affinity (K_D)	33 nM	95 nM	Catimel *et al.*, 1993 Habib *et al.*, 2013
(d) = (a)/(c)×100%	% GPA occupancy	~3%	~1%	
(e)	GPA molecules per late RBC precursor	50000		Burrill *et al.*, 2016
(f) = (e)×(d)/100	Bound Targeted EPO molecules per RBC precursor	1500	500	
(g)	Distance between antibody epitope and membrane	17–20 AA; ~60 Å	34–38 AA; ~120 Å	
(h)	Linker length	5 AA; 15 Å	35 AA; 105 Å	
(i) = (h)×2, if (h) < (g) (i) = (g) + (h), if (h) > (g)	Height spanned by Targeted EPO	30 Å	210 Å	
(j)	Surface area of a late RBC precursor cell	1000 μm^2		Burrill *et al.*, 2016
(k) = (i)×(j)	Volume spanned by Targeted EPO	3 μm^3	21 μm^3	
(l) = {(f)/(b)}/{(k)×10^{-15}}*10^9	Effective concentration of Targeted EPO	830.3 nM	39.5 nM	

Two forms of Targeted EPO, IH4-5-EPO(R150A) and 10F7-35-EPO(R150A), were used to illustrate the effect of the antibody element and the linker length on the effective concentration as well as their in vitro and in vivo activities. The affinity of the IH4 nanobody for GPA is about

3-fold higher than that of the 10F7 scFv (Table A.3), leading to about a 3-fold higher amount of bound IH4-EPO than 10F7-EPO. The volume occupied by IH4-5-EPO(R150A) is about 7-fold lower than that occupied by 10F7-35-EPO(R150A) due to its shorter linker (Figure 2.8). As a result, the local concentration of IH4-5-EPO(R150A) on an RBC precursor is predicted to be about 21-fold higher than that of 10F7-35-EPO(R150A) (Table A.3).

In TF-1 proliferation assays, IH4-5-EPO(R150A) is about 24.2-fold more potent than 10F7-35-EPO(R150A), reflecting the higher effective concentration on the target cell surface (Figure 2.3B). The comparison of Targeted EPOs containing IH4 to those with 10F7 shows that IH4 fusion proteins are generally more potent than 10F7 fusion proteins by about 10-fold, higher than the 3-fold estimated based on GPA binding affinity (Table A.1). This discrepancy may be caused by the different binding epitopes of the two antibody elements on GPA. The binding epitopes of the IH4 nanobody (epitope at $_{52}$YPPE$_{55}$) and the 10F7 scFv (epitope at about $_{34}$YAATP$_{38}$) on GPA are about 60 Å and 120 Å from the membrane, respectively (Figure 2.8). In comparison, the ligand-binding site of EPO-R is about 50 Å above the membrane. Because IH4-EPO(R150A) is in close proximity with the ligand binding site of EPO-R when it is bound to GPA, EPO-R binding would occur at a higher frequency than when 10F7-EPO(R150A) is bound to GPA, contributing to the observed higher potency. The comparison of Targeted EPOs containing the same antibody element and either 5 or 35 amino acid linker shows that the short linker is only slightly more potent than the long linker, showing about a 1.2- to 2.1-fold difference (Table A.1). A long linker allows EPO to explore a larger space with several orientations, but such a high degree of freedom comes at an entropic cost. Because the additional degrees of freedom are counteracted by the entropic cost, the effect of the linker length on EPO-R binding would be smaller than predicted. This is consistent with the previous observation that the change in linker length between 25 and 69 Å resulted in only a small change in binding energy of <2 kcal/mol [143].

In mice, reticulocyte counts were elevated by 12% on Day 4 after an i.p. injection of either 40 pmol of IH4-5-EPO(R150A) or 140 pmol of 10F7-35-EPO(R150A) (Figure 2.3C) [33]. Based on the dissociation constant (K_D) of the antibody element, 97.7% of IH4-5-EPO(R150A) and 93.0% of 10F7-35-EPO(R150A) are expected to bind mature RBCs, leaving 2.3% (0.92 pmol) and 7.0% (9.8 pmol) in the plasma, respectively. This calculation suggests that the same level of reticulocyte production was achieved when there was approximately a 10-fold lower amount of IH4-5-EPO(R150A)

available to equilibrate into the bone marrow, compared to 10F7-35-EPO(R150A). These results show that IH4-5-EPO(R150A) is about 10-fold more potent and would require about a 10-fold lower minimal effective dose than 10F7-35-EPO(R150A) in vivo. The fold difference observed in vivo, however, is less than the 21-fold predicted by the calculation of the effective concentration and the 24.2-fold measured by the TF-1 proliferation assays. This discrepancy might be explained by the previous observations that the linker length does not have a significant effect on multivalent receptor–ligand binding [143], suggesting that the avidity effect is primarily driven by the binding affinity and the position of the binding epitope of the antibody element.

As demonstrated in Chapter 2.3.3, the length of the linker in a fusion protein controls cell–cell interaction by modulating the ability of protein domains to interact with their respective receptors in trans. In contrast, varying the linker length does not seem to affect the avidity effect of a fusion protein on target cells as long as the linker is long enough to spatially allow the binding of protein domains to their respective receptors in cis on the same cell surface. By comparison, the choice of the antibody element determines both cell–cell interaction and avidity effect. In the design of a protein-based synthetic biological system, one should address questions about quantitative properties and spatial orientations of all protein domains in a quantitative manner to achieve the desired behavior.

A.5   Large Scale Production of Targeted EPO for Commercialization

After in vitro and in vivo testing of several forms of Targeted EPO, IH4-5-EPO(R150A) was chosen as the final lead molecule because of its efficacy and safety as described in Chapter 2.3. The production pipeline was optimized to produce high quality of IH4-5-EPO(R150A) that is suitable for pre-clinical studies in large quantities. The Chinese hamster ovary (CHO) cell line is a standard expression host for the production of protein therapeutics in the industry. It is particularly useful for producing EPO because glycosylation, particularly sialic acid content, is crucial for bioactivity, pharmacokinetics, and immunogenicity of EPO in vivo. Thus, recombinant EPO must be produced in cell lines, such as CHO, that mimic the natural human glycosylation patterns as much as possible. In this study, I used the CHO DG44 cell line, which was adapted to grow in serum-free suspension culture for effective protein production and engineered to lack the dihydrofolate reductase (DHFR)

gene for selection of transfected cells in a hypoxanthine/thymidine (HT)-deficient medium.

Figure A.6: *Characterization of CHO DG44 cells co-transfected with IH4-5-EPO(R150A), Cas9, and gRNA for the hamster LGMN gene. (A) Expression of IH4-5-EPO(R150A) and CRISPR knockout of LGMN shown by Western blotting. Protein expression levels were quantified by the band density and were normalized by the cell density. (B) Clone D11 has an insertion (T) at the Cas9 cleavage site, which causes a frame shift in the LGMN gene.*

Previously, one of the host cell proteins, called legumain (LGMN), was co-purified with Targeted EPO proteins in a nickel column. The crystal structure of LGMN (PDB ID: 4FGU, 4AW9, 4AWA, 4AWB) revealed that LGMN is rich in histidine residues, with patches of two to three histidine residues aligned next to one another in a folded, three-dimensional structure; this presents a His tag-like epitope to the nickel resins. LGMN was not separable from Targeted EPO by size exclusion chromatography (SEC) because its molecular weight (about 60 kDa) was too close to that of Targeted EPO (about 50 kDa). Although the presence of LGMN does not interfere with EPO activity in vitro and in vivo, it, being an asparagine endopeptidase, poses a risk of degrading Targeted EPO when the proteins are not stored properly or are exposed to conditions amenable to its proteolytic activity. In order to remove this contaminating host cell protein, I co-transfected IH4-5-EPO(R150A), a guide RNA (gRNA) targeting either exon 4 or 5 of the hamster LGMN gene (gRNA42 or 52), and Cas9. After selection in a HT-deficient medium, clones that were transfected with IH4-5-EPO(R150A) were assayed for expression and secretion of IH4-5-EPO(R150A) and LGMN in the culture medium by Western blotting. All clones but

one expressed both IH4-5-EPO(R150A) and LGMN at varying levels. Clone D11 expressed a relatively low level of EPO, and did not express any detectable level of LGMN (Figure A.6A). Sanger sequencing results indicated that only clone D11 had a frameshift mutation, an insertion (T) at the Cas9 cleavage site within exon 5 of the hamster LGMN gene, supporting the lack of expression observed by Western blotting (Figure A.6B). Unfortunately, this engineered cell line did not recover from a freeze-thaw cycle, potentially due to its lack of functional LGMN or integration of the IH4-5-EPO(R150A) gene at an unfavorable region of the host chromosome.

Figure A.7: *Generation of stable production cell lines. (A) Viability curves during MTX amplification (1 μM). MTX was removed on Day 23 post-insult to let the cells recover. (B) Protein productivity of various clones on Day 4 post-seeding in 6-well plates after the second round of MTX amplification and clonal selection. (C) Protein productivity of selected clones on Day 4 post-seeding, showing the expression levels of both Targeted EPO and LGMN from a 30-mL batch culture. Clone B9-D4 corresponds to B9-1D4 in (B). (D) Protein productivity of selected clones in a 30-mL batch culture by days. Day 0 contains some proteins that were transferred over from the previous culture. Day 7 seems to be the optimal day to harvest the cell supernatant. Clone B910 is B9-D4, and clone G810 is G8-E7 in (C).*

As an alternative to generating a LGMN knockout cell line, purification procedures were optimized to remove as much LGMN as possible from IH4-5-EPO(R150A). CHO DG44 cells transfected with IH4-5-EPO(R150A) were selected in a HT-deficient medium, and two rounds

of genomic amplification were performed to enhance the expression level of IH4-5-EPO(R150A). Clones B9, D5, and G8 were selected after the first round of genomic amplification using 100 nM of methotrexate (MTX), followed by clonal selection by limiting dilution. These three clones were expanded, and were subsequently challenged with 1 μM of MTX. The cell viability dropped significantly, and recovered as cells that were responsive to MTX survived and proliferated (Figure A.7A). After recovery, monoclonal populations were isolated by limiting dilution, and the protein productivity of each clone was assessed by Western blotting of EPO and LGMN on Day 4 post-seeding. Clones B9-D4 (or B910), D5-E10, and G8-E7 (or G810) expressed high levels of IH4-5-EPO(R150A) and low levels of LGMN (Figure A.7B–D). The viability of clone D5-E10 remained around 70%, however, implying a growth defect due to overproduction of IH4-5-EPO(R150A). An optimal time to harvest cell supernatant for protein purification was estimated by monitoring protein expression in a batch culture on Day 0, 4, 7, and 10. Day 0 contained some proteins, carried over from the previous culture during passaging, and Day 7 showed peak expression, indicating that the seventh day after seeding production cells is the optimal time to harvest cell supernatant for downstream processing of expressed proteins (Figure A.7D).

Clone B910 was batch-cultured in a 2-L volume for 7 days, and its culture supernatant was harvested. His–nickel purification was first performed to isolate IH4-5-EPO(R150A). During this process, many other impurities, including the host cell protein LGMN, were co-eluted (Figure A.8A). Because LGMN could not be separated from IH4-5-EPO(R150A) by size, I ran the elution fractions from the His–nickel purification (fractions 1 and 2 from 300 mM imidazole) through an anion exchange chromatography (AIEX) column to separate the two proteins by charge. The isoelectric points (pI) of IH4-5-EPO(R150A) and LGMN are variable depending on their sialic acid contents, and are expected to range around 4–5.5 and 2.5–4, respectively, indicating that LGMN is more negatively charged than EPO. Consistently, the more negatively charged LGMN eluted at higher salt concentrations than IH4-5-EPO(R150A) in AIEX. However, the separation was not complete, as LGMN and IH4-5-EPO(R150A) with overlapping pI (around 4) co-eluted (Figure A.8B). Earlier fractions from AIEX containing mostly IH4-5-EPO(R150A) and a little bit of LGMN were subsequently run on an SEC column for a final refinement. Despite having two different proteins in the injected sample, only a single peak was observed, which is consistent

Figure A.8: *Purification of IH4-5-EPO(R150A) from CHO DG44 clone B910. (A) His–nickel purification from the crude cell supernatant. Purification fractions were analyzed by an SDS–PAGE gel stained with Coomassie Blue. Fractions 1 and 2 from elution using 300 mM imidazole were combined and used in the subsequent purification step. (B) AIEX chromatogram showing two overlapping peaks, which indicate some separation of IH4-5-EPO(R150A) and LGMN. Pooled fractions were analyzed by an SDS–PAGE gel stained with Coomassie Blue. Pooled fractions 2–4 were combined and used for the subsequent purification step. (C) SEC chromatogram showing a single peak. A Coomassie Blue analysis shows that the later peaks mostly contain IH4-5-EPO(R150A).*

with the previous observation that LGMN and IH4-5-EPO(R150A) are not easily separable by SEC. A single peak also suggests that there was not a detectable level of protein aggregates in the sample. Several fractions within this single peak were collected and run on an SDS–PAGE gel. Coomassie Blue staining showed that the later fractions contained mostly IH4-5-EPO(R150A) with a negligible amount of LMGN. Fractions C1–C9 were pooled together, and were frozen for later use in animal studies (Figure A.8C). The final product contained >97% of IH4-5-EPO(R150A), and clone B910 showed a yield of 1.20 mg of pure IH4-5-EPO(R150A) per liter of cell culture seeded at $0.3x10^6$ cells/mL. This clone seems to produce about 4 mg of total IH4-5-EPO(R150A) per liter of cell culture, but more than half of the produced fusion proteins appear to be lost during the

purification process. Another drawback of the protein purification procedures described here is a trade-off between high purity and proper sialylation. In both AIEX and SEC, EPO with higher sialic acid content is co-eluted with LGMN because the net charge and size of the two proteins are very close. Fortunately, IH4-5-EPO(R150A) manufactured for human clinical trials and commercialization would not have the same issue of losing yield and a highly sialylated portion of IH4-5-EPO(R150A) because the His-tag would be removed from the fusion protein, thus eliminating the possibility of co-purifying LGMN. However, a new purification pipeline must be developed for a tag-less protein.

Appendix B

Supporting Information for Chapter 3

B.1 Supplementary Tables and Figures

Table B.1: *Summary of non-erythropoietic, tissue-protective EPO derivatives.*

Name	Description	Sequence	References
CEPO	Carbamylated EPO	Lysine → Homocitrulline	Leist *et al.*, 2004
ARA290	Helix B surface peptide (Cibinetide, Araim Pharmaceuticals)	UEELERALNSS (U = pyroglutamate)	Brines *et al*, 2008
JM4	AB loop	$_{28}$GCAEHCSLNENITVPDTKV$_{46}$	Yuan *et al.*, 2015
EPO(S104I)	EPO mutant Ser104Ile	Serine → Isoleucine at position 104	Gan *et al.*, 2012
NL100	Helix C	$_{100}$SGLRSLTTLLRA$_{111}$	Dmytriyeva *et al.*, 2019

Figure B.1: *Erythropoietic activities of different EPO mutants in unfused and fused forms. Typical TF-1 cell proliferation assays were performed to show the stimulation of cell proliferation by various EPO mutants. The same mutants fused to the IH4 nanobody via a 5 amino acid linker were tested for GPA-dependent activation of cell proliferation on the target cell surface. Data represent mean ± S.E.M. of three replicates.*

Figure B.2: *IH4-5-EPO(L108A) stimulates RBC production transgenic mice that express human GPA on their RBCs. Erythropoietic response to IH4-5-EPO(L108A) in mice is dose-dependent. Data represent mean ± S.E.M of five mice per dose group.*

Figure B.3 showing panels A and B. Panel A contains bar charts for "High cell density" and "Low cell density" with the following data table:

	STP (nM)	NMDA (mM)	Glutamate (mM)	CoCl$_2$ (µM)	Rotenone (nM)	H$_2$O$_2$ (µM)
Low	250	1	1	100	250	100
Medium	500	5	5	150	500	200
High	1000	10	10	250	1000	500

Figure B.3: *Optimization of experimental conditions for the in vitro neuroprotection assays. (A) Experiments were performed to find toxic agents that yield 30–40% cell death. Six toxic agents were tested for their ability to induce cell death in the SH-SY5Y cell line. They were added to cells plated at high (5.0×10^4 cells/well) or low (1.5×10^4 cells/well) seeding densities. Each agent was tested at three concentrations (low, medium, and high) within the range reported previously by other groups. The effect of NMDA and glutamate may be complicated by the change in pH. STP = staurosporine; NMDA = N-methyl-d-aspartic acid; CoCl$_2$ = cobalt chloride; H$_2$O$_2$ = hydrogen peroxide. (B) SH-SY5Y cells seeded at high density were challenged with 100 µM or 120 µM of CoCl$_2$ after 24 hr pre-treatment with varying concentrations of EPO(WT). Data represent mean $\pm$ S.E.M. of two or three replicates.*

Figure B.4: *Structural alignment models of EPO–EPO-R–CD131 heterocomplexes. (A–D) Models 1, 2, and 3 show some steric clashes between EPO and CD131. (E,F) Model 4 shows a slight steric clash between EPOs at the dodecameric interface. (C,E) Top-down views (looking down to the membrane) of heterocomplex. (A,B,D,F) Side views (perpendicular to the membrane) of a single receptor–ligand complex unit, showing EPO, EPO-R, CD131 D1, and CD131 D4. EPO residues that are within 4 Å of CD131 D1 and D4 are shown in yellow sticks. Steric clashes are indicated in red circles.*

Figure B.5: *Structural alignment models of EPO–EPO-R–CD131 heterocomplexes in which CD131 interfaces with the weak side of EPO. (A,B) Model 6 shows severe steric clashes between EPOs as well as between EPO-R and CD131 at the dodecameric interface. (C,D) Model 7 shows a steric clash between EPOs at the dodecameric interface. It also shows a slight steric clash between side chains of EPO and CD131. (A,C) Top-down views (looking down to the membrane) of heterocomplex. (B,D) Side views (perpendicular to the membrane) of a single receptor–ligand complex unit, showing EPO, EPO-R, CD131 D1, and CD131 D4. EPO residues that are within 4 Å of CD131 D1 and D4 are shown in yellow sticks. Steric clashes are indicated in red circles.*

Figure B.6: *Structural alignment models of EPO–EPO-R–CD131 heterocomplexes in which CD131 interfaces with the strong side of EPO. (A,B) Model 8 does not show any steric clash. (C,D) Model 10 shows a slight steric clash between side chains of EPO and CD131. (A,C) Top-down views (looking down to the membrane) of heterocomplex. (B,D) Side views (perpendicular to the membrane) of a single receptor–ligand complex unit, showing EPO, EPO-R, CD131 D1, and CD131 D4. EPO residues that are within 4 Å of CD131 D1 and D4 are shown in yellow sticks.*

Figure B.7: *Amino acid residues that are in contact with a neighboring receptor or ligand in alignment models 1–4, 6–8, and 10. (A,C,D–F) Models 1, 3, 4, 6, and 7 show contact areas spanning helices A and C of EPO. (B,G) Models 2 and 8 show contact areas spanning helices A and D as well as the AB loop of EPO. (H) Model 10 shows a broad contact region spanning the N and C termini of EPO. EPO, CD131 D1, and CD131 D4 are shown as dark gray, cyan, and blue cartoons, respectively. Residues that are within 4 Å of nearby receptor/ligand are shown as sticks. Among these, EPO residues are shown in yellow, with those studied in this chapter highlighted in orange.*

Appendix C

Supporting Information for Chapter 4

C.1 Sequences of Resurfaced EPO Variants

1. DNA sequences

- rsEPO-1
  ```
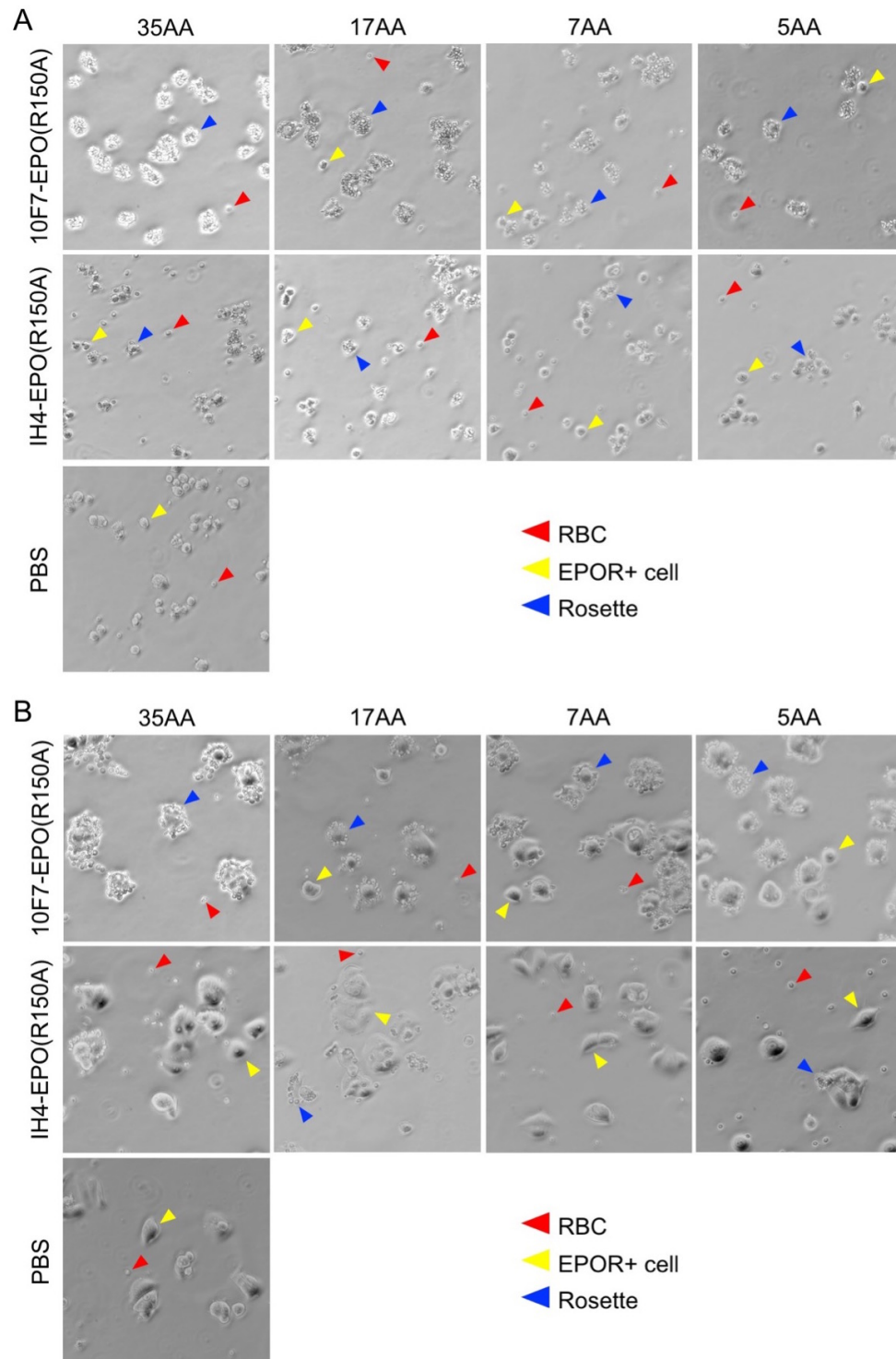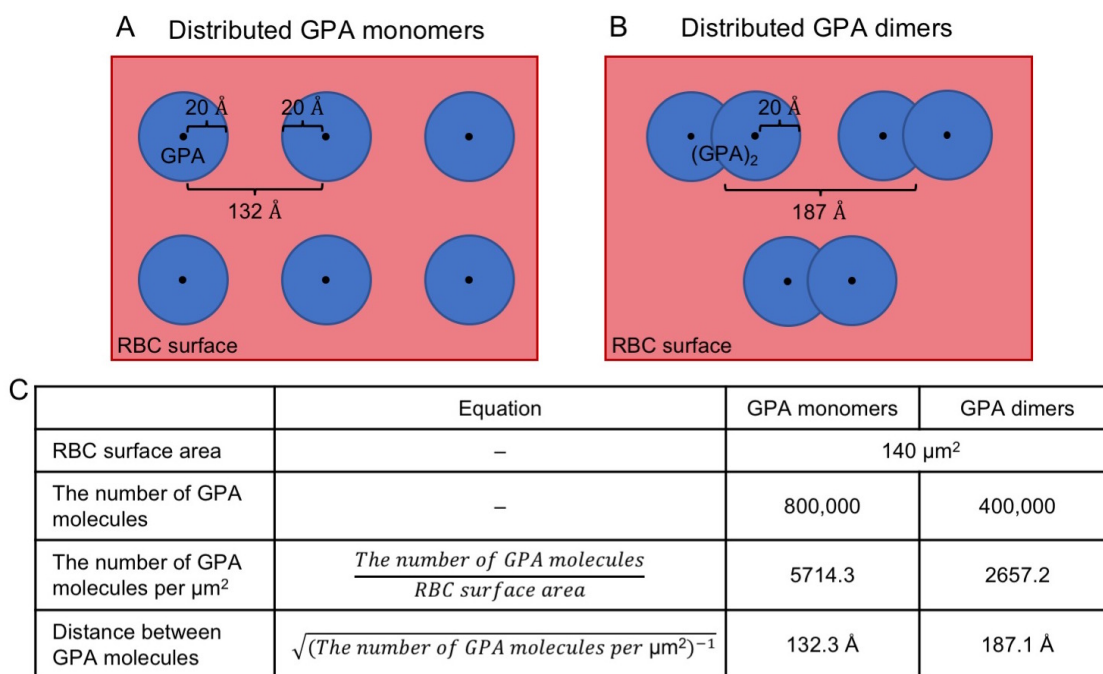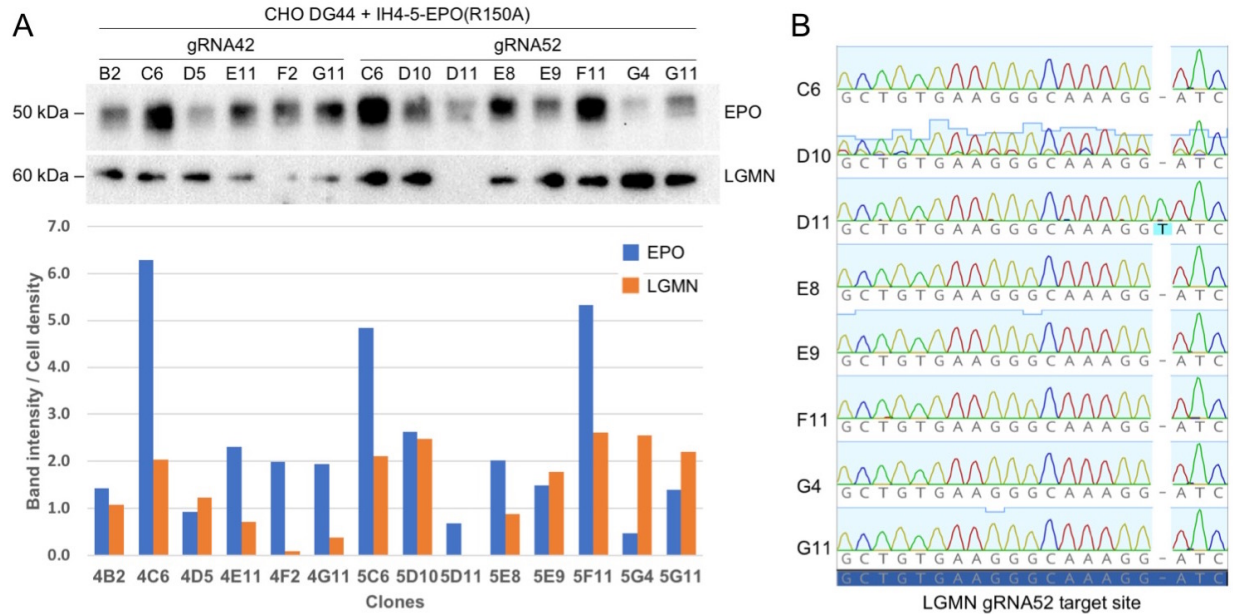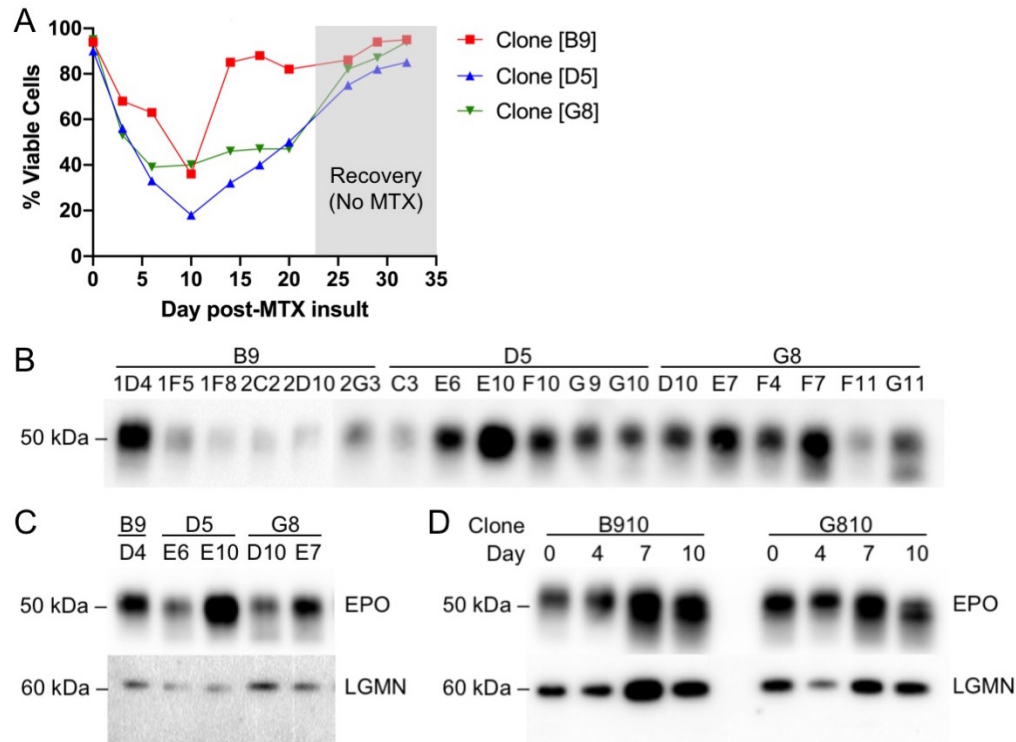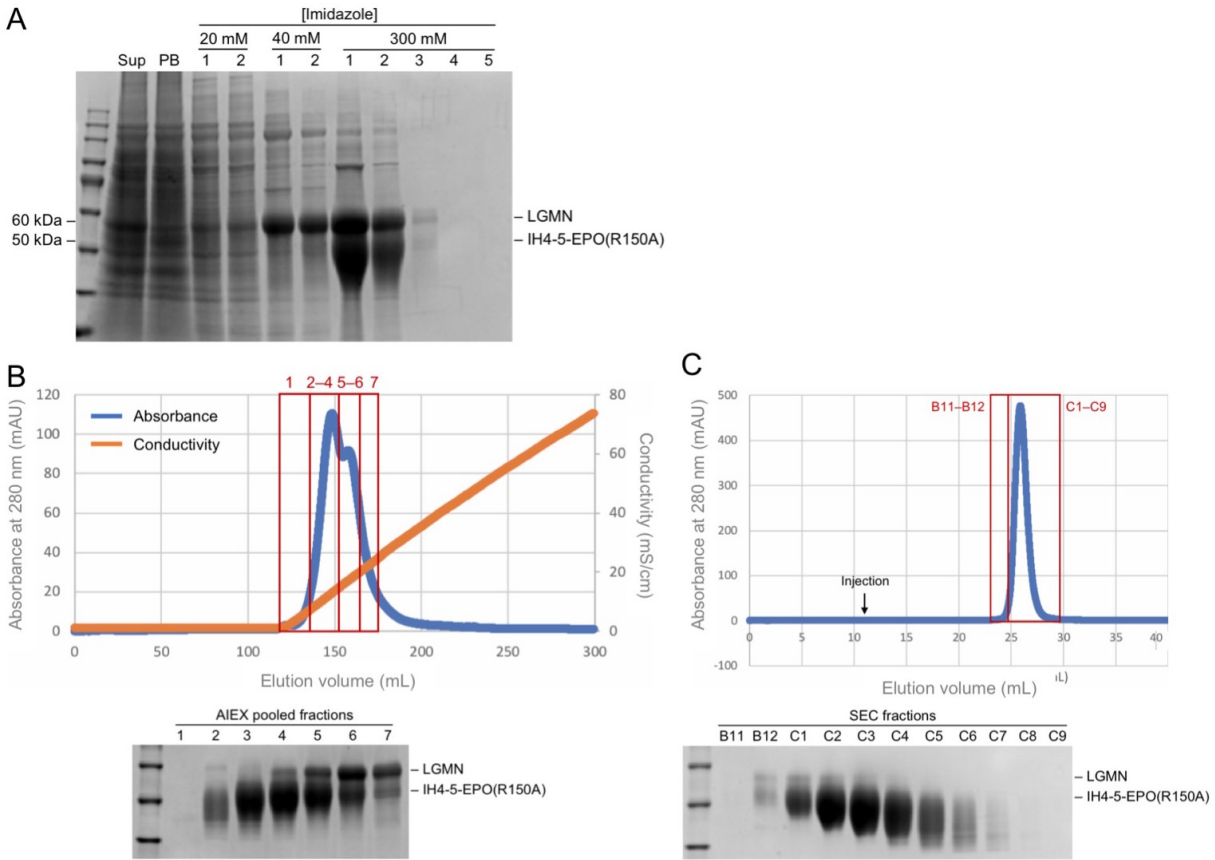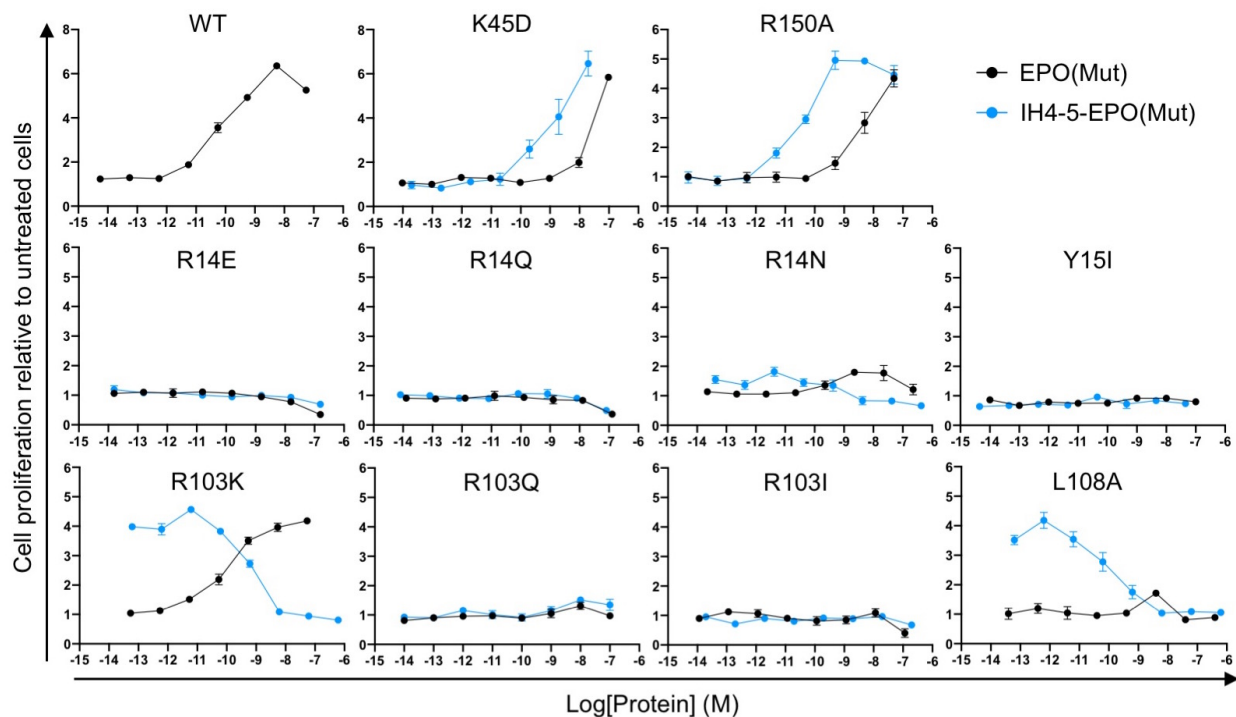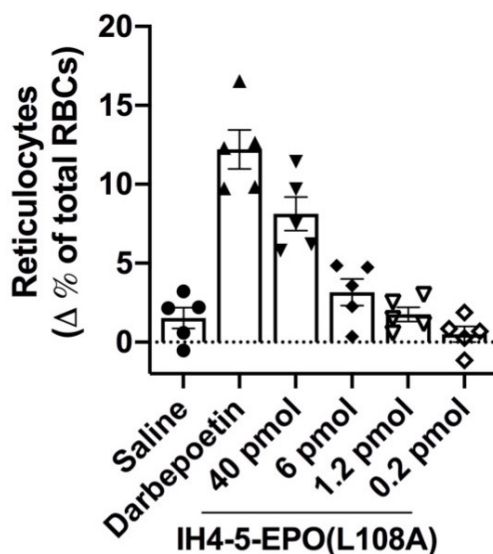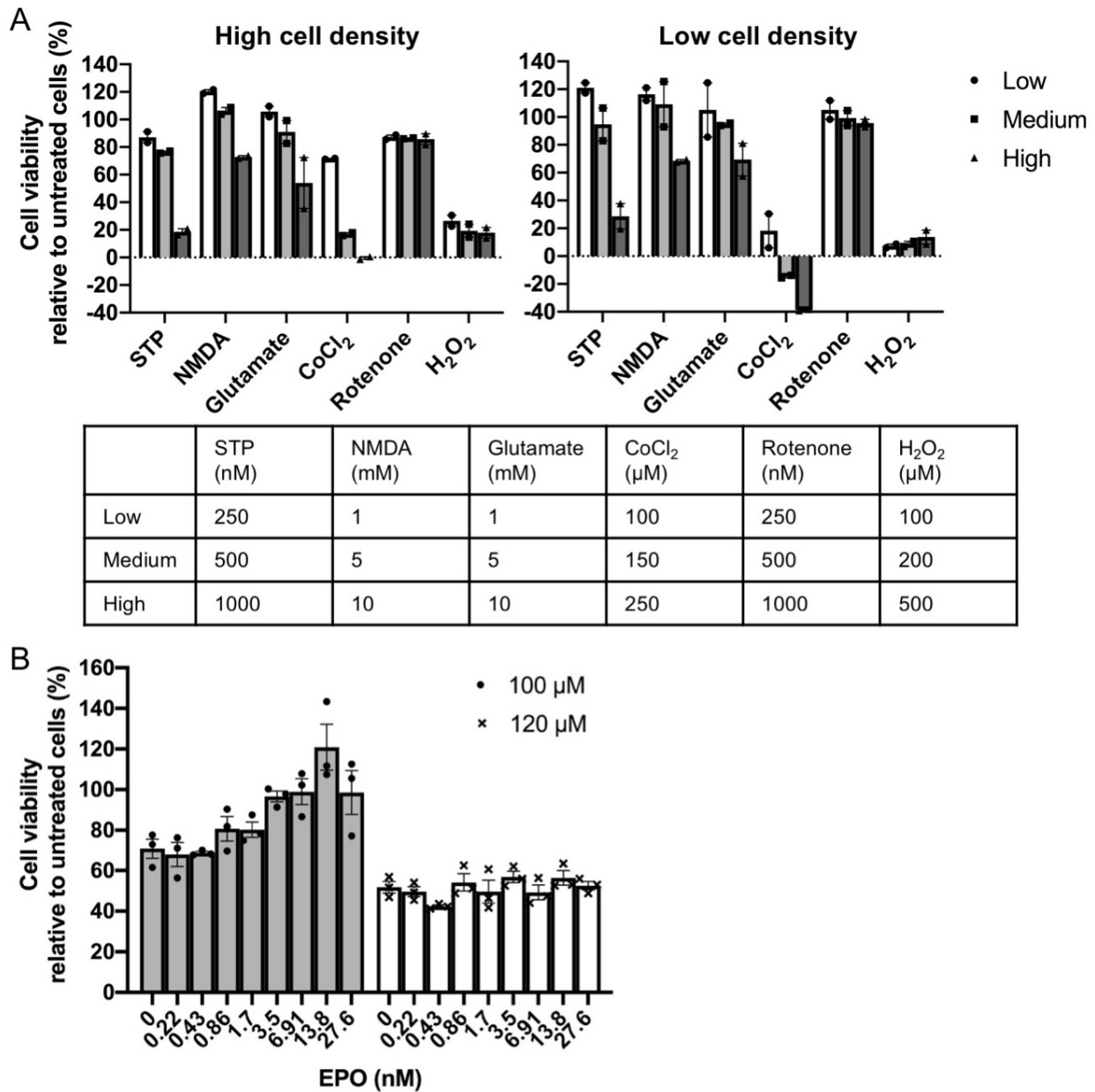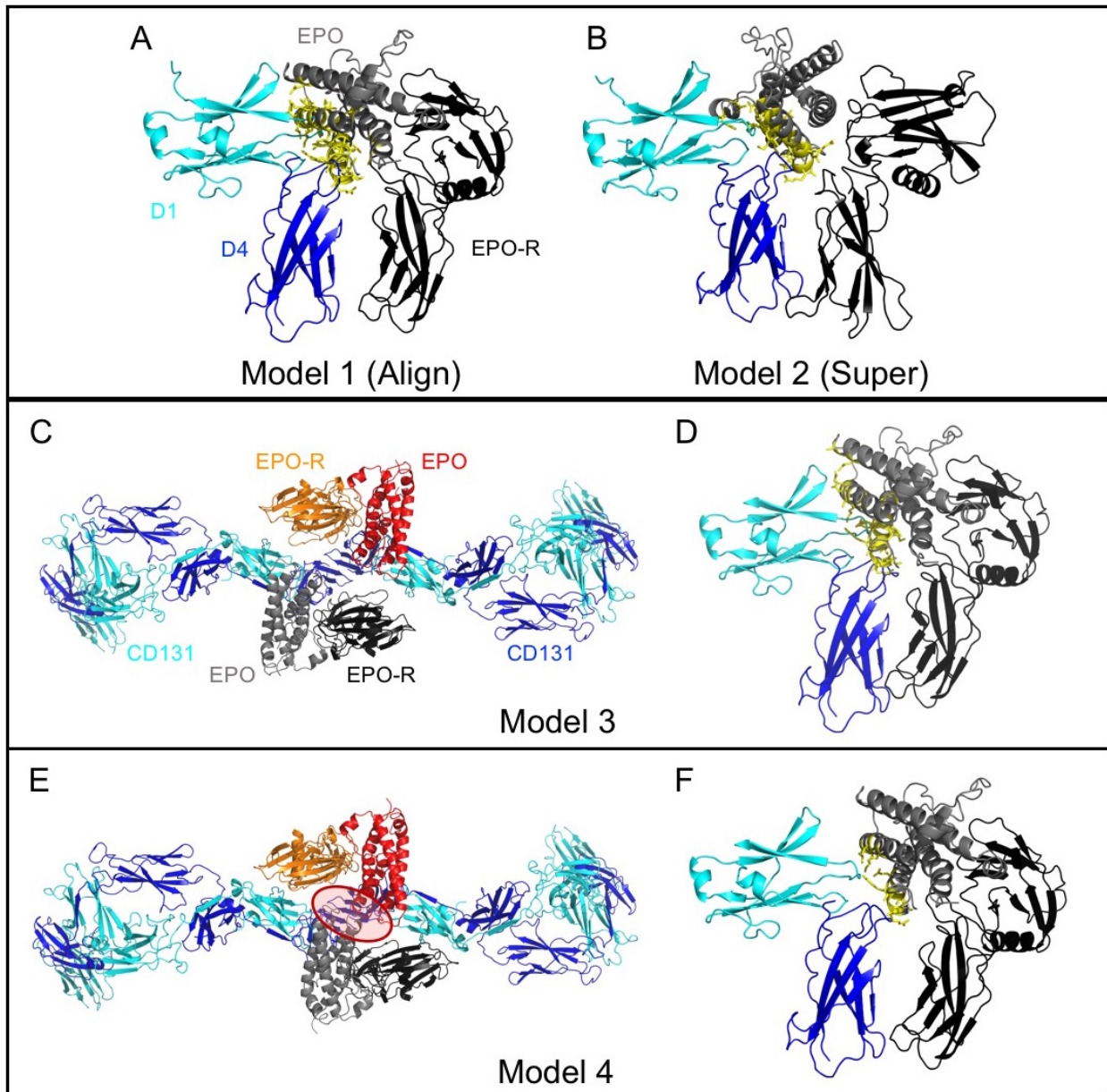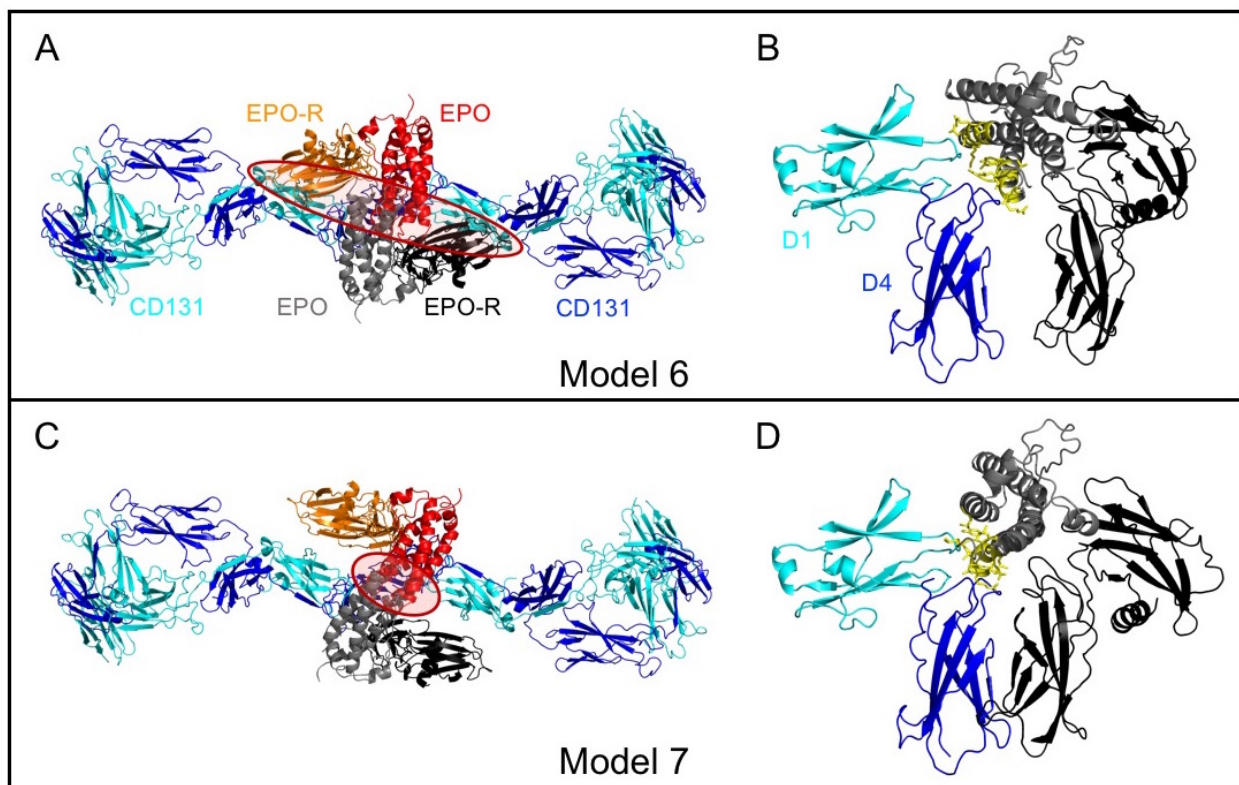  GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
  GAGGCTGAAA ATGTGACTAT GGGTTGTGCT GAAGGCTGTT CTTTCAGCGA GAATATTACT
  GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
  GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
  TTGGCCAATT CTTCTCAACC AAGTGAAACC TTGCAATTGC ATGTTGATAA GGCTATCTCT
  GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
  CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTTTCA CTGTGGATAC TCTGTGCAAG
  TTGTTTAGAA TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
  TGTAGAAGAG GTGATCGG
  ```

- rsEPO-2
  ```
  GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
  GAGGCTGAAA ATGTGACTAT GGGTTGTGCT GAAGGCTGTT CTTTCAGCGA GAATATTACT
  GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
  GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
  TTGGCCAATT CTTCTCAACC AAGTGAAACC TTGCAATTGC ATGTTGATAA GGCTATCTCT
  GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
  CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTTTCA CTGTGGATGC TCTGAGCAAG
  TTGTTTAGAA TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
  TGTAGAAGAG GTGATCGG
  ```

- rsEPO-3
  ```
  GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
  GAGGCTGAAA ATGTGACTAT GGGTTGTAAC GAAACTTGTT CTTTCAGCGA GAATATTACT
  GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
  GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
  ```

```
TTGGCCAATT CTTCTCAAGT GAACGAAACC TTGCAATTGC ATGTTGATAA GGCTATCTCT
GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTTTCA CTGTGGATGC TCTGAGCAAG
TTGTTTAGAA TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAAGAG GTGATCGG
```

- rsEPO-4

```
GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
GAGGCTGAAA ATGTGACTAT GGGTTGTGCT GAAGGCCCTT CTTTCAGCGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
TTGGCCAATG CCTCTCAACC ATGCGAAGCA TTGAGATTGC ATGTTGATAA GGCTATCTCT
GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTTTCA CTGTGGATGC TCTGAGCAAG
TTGTTTAGAA TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAAGAG GTGATCGG
```

- rsEPO-5

```
GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
GAGGCTGAAA ATGTGACTAT GGGTTGTAAC GAAACTCCTT CTTTCAGCGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
TTGGCCAATG CCTCTCAACC ATGCGAAGCA TTGAGATTGC ATGTTGATAA GGCTATCTCT
GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTGCTTCTGC CGCTCCATTG AGAACTTTCA CTGTGGATGC TCTGAGCAAG
TTGTTTAGAA TTTACTCTAA CTTCTTGAGA GGTAAGTTGA AGTTGTACAC TGGTGAAGCT
TGTAGAAGAG GTGATCGG
```

- rsEPO-6

```
GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
GAGGCTGAAA ATGTGACTAT GGGTTGTGCT GAAGGCCCTT CTTTCAGCGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
TTGGCCAATG CCTCTCAACC ATGCGAAGCA TTGAGATTGC ATGTTGATAA GGCTATCTCT
GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTTCT
CCTCCAGATG CTACTCCCTC TGCCGCTCCA TTGAGAACTT TCACTGTGGA TGCTCTGAGC
AAGTTGTTTA GAATTACTC TAACTTCTTG AGAGGTAAGT TGAAGTTGTA CACTGGTGAA
GCTTGTAGAA GAGGTGATCG G
```

- rsEPO-7

```
GCTCCACCTA GATTGATTTG TGATTCCAGA GTTTTGGAAA GATACTTGTT GGAAGCTCGG
GAGGCTGAAA ATGTGACTAT GGGTTGTGCT GAAGGCCCTT CTTTCAGCGA GAATATTACT
GTTCCAGATA CTAAGGTTAA CTTTTACGCT TGGAAGAGAA TGGAAGTTCA GCAGCAAGCT
GTTGAAGTTT GGCAAGGTTT GGCTTTGTTG TCTGAAGCTA TCTTGCAGGG TCAAGCTTTG
TTGGCCAATG CCTCTCAACC ATGCGAAGCA TTGAGATTGC ATGTTGATAA GGCTATCTCT
GGTTTGAGAT CTTTGACTTC CTTGTTGAGA GCTTTGGGTG CTCAAAAGGA AGCTATTCCC
CTGCCAGATG CTACTCCCTC TGCCGCTCCA TTGAGAACTT TCACTGTGGA TGCTCTGAGC
AAGTTGTTTA GAATTACTC TAACTTCTTG AGAGGTAAGT TGAAGTTGTA CACTGGTGAA
GCTTGTAGAA GAGGTGATCG G
```

- rsEPO-8
  ```
  ACCCCTGCTC ACCTGATCTG CGACAGTAGA GTGCTGGAGA GATACATCCT GGAGGCTAGA
  GAGGCCGAGA ACGTGACCAT GGGCTGCGCC GAGGGACCTA GCTTCAGCGA GAACATCACC
  GTGCCCGACA CCAAGGTGAA CTTCCACGCC TGGCAGAGAA TGGAGGTGCA GGAGCAGGCT
  GTGGAGGTGT GGCAGGGCCT GGCCCTGCTG AGCGAGGCTA TCCTGCAGGC CCAGGCTCTG
  CTGGCCAACA GCAGCCAGCC CTGCGAGACC CTGAGACTGC ACGTGGAGAA GGCCATCAGC
  AGCCTGAGAA GCCTGACCAG CCTGCTGAGA GTGCTGGGAG CCCAGAGAGA GGCCACCAGC
  CCTCCAGACG CCAATAGCAC CGCTCCTCTG AGAACCTTCA CCGTGGACGC CCTGAGCAAG
  CTGTTCAGAA TCTACAGCAA CTTCCTGAGA GGCAAGCTGA AGCTGTACAC CGGAGAGGCC
  TGCAGAACCG GAGACAGA
  ```

- rsEPO-9
  ```
  ACCCCTCCTA GACTGATCTG TGACAGCAGA GTTCTGGAGA GATATATTCT GGAAGCCAGA
  GAGGCTGAGA ACGTGACCAT GGGCTGCGCC GAGGGTCCTA GCTTCAGCGA GAACATCACC
  GTGCCTGACA CCAAAGTGAA TTTTCACGCC TGGAAGAACA TGAGCGTGCA GGAGCAGGCT
  GTGGAAGTGT GGCAGGGTCT GGCTCTGCTG AGCGAGGCCA TCCTGAGAGC TCAGGCTCTG
  CTGGCCAACA GCAGCCAGCC TTGTGAGACA CTGCAGCTGC ATGTGGACAA GGCTGTGAGC
  GGTCTGAGAA GCCTGACCAG CCTGCTGAGA GCCCTGGGCG TGCAGAGAGA GGCCATCAGC
  CCTCCAGACG CTGCCAGCGC TGCACCTGCC AGAACCTTTA CCGTGACCGC CCTGAGCAAG
  CTGTTCAGAA TCTACAGCAA CTTCCTGAGA GGCAAGCTGA CCCTGTACAC CGGCAGCGCC
  TGCAGAACCG GCAACAGCAC CGGCCAG
  ```

- rsEPO-10
  ```
  ACTCCTGCTC ATATCATTTG CGACGCCAGA GTGCTGGAAA GATATATCTA CGAGGCTAGA
  GAGGCTGAGA ACGTGACCAT GGGTTGCAAC GAGAGCAGCA GATTCAGCGA GAACATCACC
  GTGCCTGATA CCAAGGTGAA CTTCCACGCC TGGAAGAACA TGAGCGTGCA GGAGCAAGCT
  GTAGAAGTGT GGCAGGGTCT GGCACTGCTA AGTGAGGCCA TCCTGAGAGC TCAAGCCCTG
  CTGGCTAATA GCAGCCAGCC TAACGAGACC CTGCAGCTGC ACGTGGACAA AGCCGTGAGC
  GGTCTGAGAA GCCTGACCAG CCTGCTGAGA GCTCTGGGAG TGCAGAGAGA GGCCATCAGT
  CCACCTGACG CTGCCAGCGC TGCACCCTCC AGAACCTTCA CCGTGACAGC CCTGTGCAAG
  CTGTTCAGAA TCTACAGCAA CTTCCTGAGA GGCAAGCTGA CCCTGTACAC CGGCAGCGCC
  TGTAGAGGCG GCAATAGCAC CGGACAG
  ```

2. Protein sequences

- rsEPO-1
  ```
  APPRLICDSR VLERYLLEAR EAENVTMGCA EGCSFSENIT VPDTKVNFYA WKRMEVQQQA
  VEVWQGLALL SEAILQGQAL LANSSQPSET LQLHVDKAIS GLRSLTSLLR ALGAQKEAIS
  PPDAASAAPL RTFTVDTLCK LFRIYSNFLR GKLKLYTGEA CRRGDR
  ```

- rsEPO-2
  ```
  APPRLICDSR VLERYLLEAR EAENVTMGCA EGCSFSENIT VPDTKVNFYA WKRMEVQQQA
  VEVWQGLALL SEAILQGQAL LANSSQPSET LQLHVDKAIS GLRSLTSLLR ALGAQKEAIS
  PPDAASAAPL RTFTVDALSK LFRIYSNFLR GKLKLYTGEA CRRGDR
  ```

- rsEPO-3
  ```
  APPRLICDSR VLERYLLEAR EAENVTMGCN ETCSFSENIT VPDTKVNFYA WKRMEVQQQA
  ```

```
VEVWQGLALL SEAILQGQAL LANSSQVNET LQLHVDKAIS GLRSLTSLLR ALGAQKEAIS
PPDAASAAPL RTFTVDALSK LFRIYSNFLR GKLKLYTGEA CRRGDR
```

- rsEPO-4

```
APPRLICDSR VLERYLLEAR EAENVTMGCA EGPSFSENIT VPDTKVNFYA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANASQPCEA LRLHVDKAIS GLRSLTSLLR ALGAQKEAIS
PPDAASAAPL RTFTVDALSK LFRIYSNFLR GKLKLYTGEA CRRGDR
```

- rsEPO-5

```
APPRLICDSR VLERYLLEAR EAENVTMGCN ETPSFSENIT VPDTKVNFYA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANASQPCEA LRLHVDKAIS GLRSLTSLLR ALGAQKEAIS
PPDAASAAPL RTFTVDALSK LFRIYSNFLR GKLKLYTGEA CRRGDR
```

- rsEPO-6

```
APPRLICDSR VLERYLLEAR EAENVTMGCA EGPSFSENIT VPDTKVNFYA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANASQPCEA LRLHVDKAIS GLRSLTSLLR ALGAQKEAIS
PPDATPSAAP LRTFTVDALS KLFRIYSNFL RGKLKLYTGE ACRRGDR
```

- rsEPO-7

```
APPRLICDSR VLERYLLEAR EAENVTMGCA EGPSFSENIT VPDTKVNFYA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANASQPCEA LRLHVDKAIS GLRSLTSLLR ALGAQKEAIP
LPDATPSAAP LRTFTVDALS KLFRIYSNFL RGKLKLYTGE ACRRGDR
```

- rsEPO-8

```
TPAHLICDSR VLERYILEAR EAENVTMGCA EGPSFSENIT VPDTKVNFHA WQRMEVQEQA
VEVWQGLALL SEAILQAQAL LANSSQPCET LRLHVEKAIS SLRSLTSLLR VLGAQREATS
PPDANSTAPL RTFTVDALSK LFRIYSNFLR GKLKLYTGEA CRTGDR
```

- rsEPO-9

```
TPPRLICDSR VLERYILEAR EAENVTMGCA EGPSFSENIT VPDTKVNFHA WKNMSVQEQA
VEVWQGLALL SEAILRAQAL LANSSQPCET LQLHVDKAVS GLRSLTSLLR ALGVQREAIS
PPDAASAAPA RTFTVTALSK LFRIYSNFLR GKLTLYTGSA CRTGNSTGQ
```

- rsEPO-10

```
TPAHIICDAR VLERYIYEAR EAENVTMGCN ESSRFSENIT VPDTKVNFHA WKNMSVQEQA
VEVWQGLALL SEAILRAQAL LANSSQPNET LQLHVDKAVS GLRSLTSLLR ALGVQREAIS
PPDAASAAPS RTFTVTALCK LFRIYSNFLR GKLTLYTGSA CRGGNSTGQ
```

C.2   Protein Sequences of Mammalian Wild-Type EPO

- Human (Homo sapiens; GenBank CAA26095.1)

```
APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA
VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS
PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDR
```

- Rhesus monkey (Macaca mulatta; GenBank AAA36842.1)

```
APPRLVCDSR VLERYLLEAK EAENVTMGCS ESCSLNENIT VPDTKVNFYA WKRIEVGQQA
```

```
VEVWQGLALL SEAVLRGQAV LANSSQPFEP LQLHMDKAIS GLRSITTLLR ALGAQEAISL
PDAASAAPLR TITADTFCKL FRVYSNFLRG KLKLYTGEAC RRGDR
```

- Mouse (Mus musculus; GenBank AAI44884.1)
```
APPRLICDSR VLERYILEAK EAENVTMGCA EGPRLSENIT VPDTKVNFYA WKRMEVEEQA
IEVWQGLSLL SEAILQAQAL LANSSQPPET LQLHIDKAIS GLRSLTSLLR VLGAQKELMS
PPDTTPPAPL RTLTVDTFCK LFRVYANFLR GKLKLYTGEV CRRGDR
```

- Rat (Rattus norvegicus; NCBI NP_058697.1)
```
APPRLICDSR VLERYILEAK EAENVTMGCA EGPRLSENIT VPDTKVNFYA WKRMKVEEQA
VEVWQGLSLL SEAILQAQAL QANSSQPPES LQLHIDKAIS GLRSLTSLLR VLGAQKELMS
PPDATQAAPL RTLTADTFCK LFRVYSNFLR GKLKLYTGEA CRRGDR
```

- Damara mole-rat (Fukomys damarensis; NCBI XP_010607915.1)
```
VPQSLICDSR VLERYILEAK DAENVTMGCM EHCSLGENIT VPDTKVNFYA WQRMEVGQQA
AEVWQGLVLL SEAILRGQAL LANSSQPSET LRLHMDKAVS GLRSLTSLLR TLGAQKEPIL
PPDAASPAPL RTFTVDALCK LFRVYSNFLR GKLKLYTGEA CRRGDR
```

- Prairie vole (Microtus ochrogaster; NCBI XP_013210725.1)
```
APPRLICDSR VLERYILEAR EAENVTMGCA EGPRLSENIT VPDTKVNFNA WKRMEVQEQA
VEVWQGLSLL SEAILRGQAL LANSSQPSGM LQLHIDKAIS GLRSLTSLLR VLGAQKESLS
PPDATPPAPL RTLMVENFCK LFRVYSNFLR GKLKLYTGEA CRRGDR
```

- Chinese hamster (Cricetulus griseus; GenBank EGW06331.1)
```
APPRLICDSR VLERYILEAK EAENVTMGCA EGPRLSENIT VPDTKVNFYA WKRKAVEEQA
VEVWHGLSLL SEAILRGQAL LGNSSQPSEM LQLHIDKAIS GLRSLTSLLR VLGAQKESIL
PPDATPPAPL RTLSVDTFCK LFRVYSNFLR GKLKLYTGEA CRRGAR
```

- Thirteen-lined ground squirrel (Ictidomys tridecemlineatus; NCBI XP_021584302)
```
TPARLICDSR VLERYILEAK EAENVTMGCA DGCSLSENIT VPDTKVNFYA WRRMEVGQQA
VEVWQGLALL SEAILRSQAL LANSSQPSDT LRLHVDKAVS GLRSLTSLLR AMGVQREAVL
PPDAASAAPL RTFTVDTLCK LFRIYSNFLR GKLKLYTGEA CRRGDR
```

- Great roundleaf bat (Hipposideros armiger; NCBI XP_019485449)
```
APPNLICDSR VLERYILEAK EAENVTMGCE EGCSFSENIT VPDTKVNFYA WKRMEVEQQA
VEVWQGLALL SEAIPRGQAL LANSSQPSET LKLHMDKAIS GLRSLTSLLR ALGAQKEAIS
LPDATASPAA LRTFTVDTLC KLFRIYSNFL RGKLKLYTGE ACRRGDR
```

- Rabbit (Oryctolagus cuniculus; NCBI NP_001075559.1)
```
APARLICDSR VLERYILEAK EAENVTMGCA EGCSLGENIT VPDTKVNFHH WKKSEAGRHA
VEVWQGLALL SEAMLRSQAL LANSSQLPET LQVHVDKAVS GLRSLTSLLR ALGVQKEAVS
PPEAASSAAP LRTVAADTLC KLFRIYSNFL RGKLKLYTGE ACRRGDR
```

- Pig (Sus scrofa; GenBank CAB96417.1)
```
APPRLICDSR VLERYILEAK EGENATMGCA ESCSFSENIT VPDTKVNFYA WKRMEVQQQA
MEVWQGLALL SEAILQGQAL LANSSQPSEA LQLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDASPSSAT PLRTFAVDTL CKLFRNYSNF LRGKLTLYTG EACRRRDR
```

- Red deer (Cervus elaphus hippelaphus; NCBI OWK11126.1)
```
APPRLICDSR VLERYILEAR EAENATMGCV EGCSFNENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLAQL SEAILRGQVL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDATPSAAP LRTFTVDALS KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Tibetan antelope (Pantholops hodgsonii; NCBI XP_005956299.1)

```
APPRLICDSR VLERYILEAR EAENATMGCA EGCSFNENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLALL SEAILRGQAL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDATPSAAP LRAFTVDALS KLFQIYSNFL RGKLTLYTGE ACRRGDR
```

- Sheep (Ovis aries; GenBank CAA80848.1)

```
APPRLICDSR VLERYILEAR EAENATMGCA EGCSFSENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLALL SEAIFRGQAL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDATPSAAP LRIFTVDALS KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Cow (Bos taurus; GenBank ABN48309.1)

```
APARLICDSR VLERYILEAR EAENATMGCA EGCSFNENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLALL SEAILRGQAL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIS
LPDATPSAAP LRAFTVDALS KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Yak (Bos grunniens; GenBank ABN59378.1)

```
APARLICDSR VLERYILEAR EAENATMGCA EGCSFNENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLALL SEAILRGQAL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIS
LPDATPSAAP LRAFTVDALS KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Goat (Capra hircus; NCBI XP_005697875.1)

```
APQRLICDSR VLERYILEAR EAENATMGCA EGCSFNENIT VPDTKVNFYA WKRMEVQQQA
LEVWQGLALL SEAILRGQAL LANASQPCEA LRLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDATPSAAP LRIFTVDALS KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Vicuna (Vicugna pacos; NCBI XP_006219204.1)

```
APPRLICDSR VLERYILEAR EAENATMGCA EGCSFSENIT VPDTKVNFYA WKRMEVAQQA
VEVWQGLALL SEAILQGQAL LANSSQPSEV LQLHVDKAVS GLRSLTSLLR ALGAQKEAIP
LPDAASSSAA PLRTFTVDTL CKLFRIYSNF LRGKLTLYTG EACRRGDR
```

- Baleen whale (Balaenoptera acutorostrata scammoni; NCBI XP_007186678.1)

```
APPRLICDSR VLERYILEAR EAENATMGCA EGCSFSENIT VPDTKVNFHA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANSSQPSEA LQLHVDKAVS GLRSLTSLLR ALGAQKEAVP
LPDAASSAAP LRTFTIDTLC KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Beluga whale (Delphinapterus leucas; NCBI XP_022444347.1)

```
APPHLICDSR VLERYILEAR EAENATVGCA EGCSFSENIT VPDTKVNFYA WKRMEVQQQA
VEVWQGLALL SEAILQGQAL LANSSQPSEA LQLHVDKAVS GLRSLTSLLR ALGAQKGAVP
LPDAASSAAP LRTFTVDTLC KLFRIYSNFL RGKLTLYTGE ACRRGDR
```

- Dog (Canis lupus familiaris; GenBank AAS77874.1)

```
APPRLICDSR VLERYILEAR EAENVTMGCA QGCSFSENIT VPDTKVNFYT WKRMDVGQQA
LEVWQGLALL SEAILRGQAL LANASQPSET PQLHVDKAVS SLRSLTSLLR ALGAQKEAMS
LPEEASPAPL RTFTVDTLCK LFRIYSNFLR GKLTLYTGEA CRRGDR
```

- Cat (Felis catus; GenBank AFN85670.1)

```
APPRLICDSR VLERYILEAR EAENVTMGCA EGCSFSENIT VPDTKVNFYT WKRMDVGQQA
VEVWQGLALL SEAILRGQAL LANSSQPSET LQLHVDKAVS SLRSLTSLLR ALGAQKEATS
LPEATSAAPL RTFTVDTLCK LFRIYSNFLR GKLTLYTGEA CRRGDR
```

- Horse (Equus caballus; GenBank BAC55239.1)
  ```
  APPRLICDSR VLERYILEAR EAENVTMGCA EGCSFGENVT VPDTKVNFYS WKRMEVEQQA
  VEVWQGLALL SEAILQGQAL LANSSQPSET LRLHVDKAVS SLRSLTSLLR ALGAQKEAIS
  PPDAASAAPL RTFAVDTLCK LFRIYSNFLR GKLKLYTGEA CRRGDR
  ```

- Rhino (Ceratotherium simum; NCBI XP_004442940.1)
  ```
  APPRLICDSR VLERYILEAR EAENVTMGCA EGGSFGENIT VPDTKVNFYT WKRMEVEQQA
  VEVRQGLALL SEAILQGQAL LVNSSQLSET LRLHVDKAVS SLRSLTSLFR ALGAQKESIS
  LPDTASAAPL RTFTVDTLCK LFRIYSNFLR GKLKLYTGEA CRKGDR
  ```

C.3 Supplementary Figure

Figure C.1: *Immunotoxicology studies of Targeted EPO, IH4-5-EPO(R150A), in mice. (A) Experimental design. Mice received repeated i.p. injections of darbepoetin (2 µg; 50 pmol) + BSA (50 µg), IH4-5-EPO(R150A) (2 µg; 40 pmol) + BSA (50 µg), or vehicle twice a week for four weeks. A small amount of blood was drawn by tail-nick to monitor reticulocyte levels in these mice. Dosing was stopped when the reticulocyte counts dropped below the Day 0 baseline (Day 25). (B) Reticulocyte counts were measured by flow cytometry on indicated days. Data represent mean ± S.E.M. of five mice per dose group. Some images were created with BioRender.com.*

A

```
            1          11         21         31         41         51
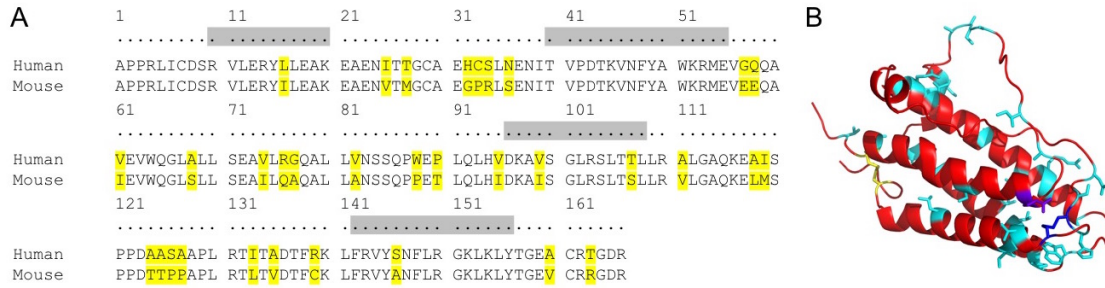        .......... .......... .......... .......... .......... ..........
Human   APPRLICDSR VLERYLLEAK EAENITTGCA EHCSLNENIT VPDTKVNFYA WKRMEVGQQA
Mouse   APPRLICDSR VLERYILEAK EAENVTMGCA EGPRLSENIT VPDTKVNFYA WKRMEVEEQA
            61         71         81         91         101        111
        .......... .......... .......... .......... .......... ..........
Human   VEVWQGLALL SEAVLRGQAL LVNSSQPWEP LQLHVDKAVS GLRSLTTLLR ALGAQKEAIS
Mouse   IEVWQGLSLL SEAILQAQAL LANSSQPPET LQLHIDKAIS GLRSLTSLLR VLGAQKELMS
            121        131        141        151        161
        .......... .......... .......... .......... ......
Human   PPDAASAAPL RTITADTFRK LFRVYSNFLR GKLKLYTGEA CRTGDR
Mouse   PPDTTPPAPL RTLTVDTFCK LFRVYANFLR GKLKLYTGEV CRRGDR
```

B

Figure C.2: *Comparison of protein sequences and structures of human and mouse EPOs. (A) Pairwise alignment of human and mouse EPO protein sequences. Two EPOs show about 80% sequence homology. Highlighted amino acids indicate differences between the two aligned sequences. Gray regions are important for EPO–EPO-R interaction and activity. (B) Structure of human EPO (PDB ID: 1EER). Residues that are different from mouse EPO are shown as cyan sticks. Disulfide bond (Cys7–Cys161) common to both species is shown in yellow. Disulfide bond (Cys29–Cys33) found in human but not in mouse is shown in blue. Residue at position 139 (Arg139), shown in purple, is a cysteine in mouse EPO and forms an alternative disulfide bond with Cys29.*

A

EPO	Disulfide bonds	Glycosylation sites	Variations in CD loop	Variations at C-terminus	Number of modifications
Mouse wildtype	C7-C161 C29-C139	N24, N38, N83			
rsEPO-1	C7-C161 **C29-C33/139**	N24, N38, N83, **S126**			24
rsEPO-2	C7-C161 **C29-C33**	N24, N38, N83, **S126**			26
rsEPO-3	C7-C161 **C29-C33**	N24, N38, N83, **S126,** **N30, N88**			29
rsEPO-4	C7-C161 **C29-C88**	N24, N38, N83, **S126**			28
rsEPO-5	C7-C161 **C29-C88**	N24, N38, N83, **S126,** **N30**			30
rsEPO-6	C7-C161 **C29-C88**	N24, N38, N83, **S126**	**P127SA**		28
rsEPO-7	C7-C161 **C29-C88**	N24, N38, N83, **S126**	**SP120PL, P127SA**		30
rsEPO-8	C7-C161 **C29-C88**	N24, N38, N83, **S126,** **N125**	**K116R, I119T,** **T125N, A127T**		31
rsEPO-9	C7-C161 **C29-C88**	N24, N38, N83, **S126,** **N53, N165**	**K116R, L130A**	**DR165NSTGQ**	40
rsEPO-10	C7-C161 C29-C139	N24, N38, N83, **S126,** **N30, N53, N88, N165**	**K116R, L130S**	**DR165NSTGQ**	46

*Amino acid numbering is based on human wildtype EPO without adding the insertions.

B

Figure C.3: *(A) A list of rsEPO variants and modifications (bold) relative to mouse wild-type EPO. (B) Pairwise protein sequence alignment of rsEPO variants to mouse wild-type EPO. Highlighted amino acids indicate the differences between the variants and mouse EPO. Boxed regions are important for EPO–EPO-R interaction and activity. Sequences can be viewed in Appendix C.*