

Charles University

Faculty of Science

Study program: Biochemistry



Ing. PHAM NGOC PHUONG

Specificity of Protein-Protein Interactions and Their Modulation

Specificita Interakcí Protein-Protein A Jejich Modulace

Doctoral thesis

Supervisor:

prof. Ing. Bohdan Schneider, CSc., DSc

Prague, 2024

Declaration

I hereby state that I have completed this thesis as an independent work under the guidance of my supervisor, prof. Ing. Bohdan Schneider, CSc., DSc. I have properly cited all literature, publication and other information sources that I have used. This thesis has not been submitted previously anywhere to obtain the same or any other academic degree.

In Prague, March 27th, 2024.

Pham Ngoc Phuong

Acknowledgments

I would like to express my heartfelt gratitude to my supervisor, professor Bohdan Schneider, who made my PhD journey come true. I am very grateful for his scientific advice, financial support and the opportunity he offered me to learn new technologies. Here, I would like to say: “Dear Bohdan, you are not only the greatest supervisor I ever had but you also inspired the scientist I want to become”.

I also would like to specially thank Dr. Gustavo Fuertes Vives for advising me the experimental performance and trouble shooting. I truly appreciate his patient and willingness for sharing his extensive knowledge and keep motivating me, and remind me to never give up and complete my PhD thesis.

I certainly would like to say many thanks to my colleagues for all their experimental help in the laboratory, and especially my good friend Dr. Pavel Mikulecký. I thank him for teaching me new methods when I first came to the lab, for taking care of my work progress and for sharing his valuable experiences about trial-and-error experiments.

I am deeply grateful to my family: my parents, my brother in Vietnam, my husband and my daughter Ana Phuong Fuertes Pham for always taking care of me, for encouraging me and for being my main source of motivation.

Phạm Ngọc Phương

Abstract (EN)

Protein-protein interactions (PPI) have essential roles in life processes, and abnormal PPI are associated with many human diseases. Given their importance, PPI have received increasing attention and became drug targets. However, the design of specific PPI and their modulation is challenging. Cytokine-receptor interactions are especially important in the regulation of the immune system. Interleukin-10 (IL-10) over-production results in excessive immunosuppressive effects, tumor growth and infection. The interaction between interferon gamma receptor 2 (IFN- γ R2) and interferon gamma (IFN- γ) leads to activation of downstream signaling pathways but the mechanism of such interaction is elusive. Interleukin-24 (IL-24) is another cytokine that signals through receptors sharing the interleukin-20 receptor 2 (IL-20R2) subunit and has important roles in autoimmunity and cancer.

The aims of this Ph.D. thesis are to study PPI from several aspects emphasizing their specificity. The first goal is to develop a novel protein scaffold and subsequently evolve it into a high-affinity binder *specific* for human IL-10. The second goal is to understand the structural basis for receptor *specificity* of human IFN- γ . The third goal is to *modulate* the binding affinity between human IL-24 and its receptor IL-20R2 by using photo-responsive non-canonical amino acids and light.

The N-terminal domain of a monomeric human protein (PIH1 domain-containing protein 1), with a fold different from previously known non-antibody scaffolds, was designed as our novel scaffold called *57aBi*. The functionality of such a new scaffold was demonstrated by training it as a nanomolar-affinity binder against IL-10 using methods of directed evolution. The structures of two binders solved by X-ray crystallography showed that the evolved proteins share a similar fold as the parental scaffold. In addition, the crystal structure of IFN- γ R2 revealed the importance of certain residues, glycosylation and disulfide bond formation on specific interactions with both interferon gamma receptor 1 (IFN- γ R1) and IFN- γ . Regarding the photocontrol of the binding between IL-24 and IL-20R2 by genetic code expansion, a photocaged tyrosine residue, nitrobenzyl-tyrosine (NBY), on IL-20R2 was found to be sufficient to diminish PPI, and binding was restored upon UV light irradiation. Another NBY substitution on IL-24 triggered the Janus Kinase/Signal Transducer and Activator of Transcription (JAK/STAT) signaling cascade upon exposure to 365-nm light.

In summary, by a combination of directed evolution, structural biology, and photoprotein engineering, I have shed light into the specificity and modulation of PPI involving cytokines and their receptors.

Abstrakt (CZ)

Protein-proteinové interakce (PPI) mají zásadní roli v životních procesech a jejich abnormality jsou často spojené s vznikem nemocí. Vzhledem ke své důležitosti dostávají PPI čím dál tím více pozornosti a staly se cíli zásahu různých léčiv. Návrh konkrétních PPI a jejich modulace je však náročný. Interakce cytokinů s jejich receptory jsou zvláště důležité při regulaci imunitního systému. Nadměrná produkce interleukinu-10 (IL-10) má za následek nadměrné imunosupresivní účinky, růst nádoru a infekci. Interakce mezi receptorem interferonu gama 2 (IFN- γ R2) a interferonem gama (IFN- γ) vede k aktivaci signálních drah, ale mechanismus takové interakce je stále nepopsaný. Interleukin-24 (IL-24) je další cytokin, který signalizuje prostřednictvím receptorů sdílejících receptorovou podjednotku 2 interleukinu-20 (IL-20R2) a má důležitou roli v autoimunitě a rakovině.

Cílem této Ph.D. práce je studium PPI z několika hledisek s důrazem na jejich specifičnost. Prvním cílem je vyvinout nové proteinové lešení, které bude následně použito pro vývoj vysokoafinitního pojiva specifického pro lidský IL-10. Druhým cílem je pochopení strukturního základu receptorové specifičnosti lidského IFN- γ . Třetím cílem je modulace vazebné afinity mezi lidským IL-24 a jeho receptorem IL-20R2 za použití foto-responzivních nekanonických aminokyselin a světla.

N-terminální doména monomerního lidského proteinu nazvaného PIH1, s odlišnou strukturou od dříve známých neprotilátkových proteinových lešení, byla navržena jako naše nové proteinové lešení nazvané 57aBi. Funkčnost této molekuly byla demonstrována jejím vývojem jako vazebného proteinu cíleného proti IL-10 s nanomolární vazebnou afinitou pomocí metod řízené evoluce. Struktury dvou vazebných látek řešené rentgenovou krystalografií ukázaly, že námi vyvinuté proteiny sdílejí podobné složení s původní molekulou. Krystalová struktura IFN- γ R2 navíc odhalila důležitost určitých aminokyselin, glykosylace a tvorby disulfidových vazeb pro specifické interakce jak s interferonovým gama receptorem 1 (IFN- γ R1), tak s IFN- γ . Pokud jde o kontrolu vazby mezi IL-24 a IL-20R2 za použití světla a expanze genetického kódu za účelem vložení nekanonické na světlo citivé aminokyseliny, bylo zjištěno, že aminokyselina nesoucí fotoklec, nitrobenzyl-tyrosin (NBY), na IL-20R2 je dostatečně účinná k potlačení vzniku PPI. Po opětovném UV ozáření došlo k obnově vazby. Pomocí další implementace NBY do IL-24 a po vystavení záření o vlnové délce 365 nm se navíc podařilo i spustit signální kaskádu JAK/STAT.

Kombinací řízené evoluce, strukturní biologie a fotoxenoproteinového inženýrství byla osvětlena problematika specifičnosti a modulace PPI zahrnující cytokiny a jejich receptory.

Publications and my contribution

1. **Pham, P. N.**, M. Huličiak, L. Biedermannová, J. Černý, T. Charnavets, G. Fuertes, Š. Herynek, L. Kolářová, P. Kolenko, J. Pavlíček, J. Zahradník, P. Mikulecky and B. Schneider (2021). "[Protein Binder \(ProBi\) as a New Class of Structurally Robust Non-Antibody Protein Scaffold for Directed Evolution.](#)" *Viruses* 13(2).

My contribution: searching structures on PDB, sequencing analysis, cloning-expression-purification of wild-type scaffold candidates and binder variants, ribosome display experiments, affinity measurement, protein crystallization.

Supplemented paper:

Kolenko, P., P. Mikulecky, **P. N. Pham**, M. Maly and B. Schneider (2023). "[Diffraction anisotropy and paired refinement: crystal structure of H33, a protein binder to interleukin 10.](#)" *J Appl Crystallogr* 56(Pt 4): 1261-1266.

My contribution: protein production and crystallization.

2. **Pham, P. N.**, J. Zahradnik, L. Kolarova, B. Schneider and G. Fuertes (2023). "[Regulation of IL-24/IL-20R2 complex formation using photocaged tyrosines and UV light.](#)" *Front Mol Biosci* 10: 1214235.

My contribution: photoxenoprotein design, cloning, expression and purification of parental IL, receptor and their nCAA variants, MST affinity measurement, cell signaling performance and protein detection by western blot.

3. Mikulecky, P., J. Zahradnik, P. Kolenko, J. Cerny, T. Charnavets, L. Kolarova, I. Necasova, **P. N. Pham** and B. Schneider (2016). "[Crystal structure of human interferon-gamma receptor 2 reveals the structural basis for receptor specificity.](#)" *Acta Crystallogr D Struct Biol* 72(Pt 9): 1017-1025.

My contribution: assistance in protein crystallization and structure determination.

List of abbreviations

Acronym or symbol	Description
aaRS	Aminoacyl-tRNA synthetase
ARPs	Ankyrin repeat proteins
APC	Antigen presenting cell
Fab	Antigen-binding fragment
AzoF	Azobenzene-phenylalanine
BSA	Bovine serum albumin
CD	Circular dichroism
CDRs	Complementarity-determining regions
Fc	Constant fragment
CSIF	Cytokine synthesis inhibitory factor
DSF	Differential scanning fluorimetry
DLS	Dynamic light scattering
ELISA	Enzyme-linked immunosorbent assays
EGFR	Epidermal growth factor receptor
EBV	Epstein-Barr virus
$\Delta\Delta G$	Free energy differences
GCE	Genetic code expansion
ABK	Azibutyl-lysine (Photo-lysine)
Igs	Immunoglobulins
IBD	Inflammatory bowel disease
IFN- γ	Interferon gamma
IFN- γ R1	Interferon gamma receptor 1
IFN- γ R2	Interferon gamma receptor 2
IFN- λ	Interferon lambda
IFN- λ R1	Interferon lambda receptor 1
IFNs	Interferons
IL-10	Interleukin-10
IL-10R1	Interleukin-10 receptor 1
IL-10R2	Interleukin-10 receptor 2
IL-19	Interleukin-19
IL-20	Interleukin-20
IL-20R1	Interleukin-20 receptor 1
IL-20R2	Interleukin-20 receptor 2
IL-22	Interleukin-22
IL-22R1	Interleukin-22 receptor 1
IL-24	Interleukin-24
IL-26	Interleukin-26
IL-28	Interleukin-28
IL-29	Interleukin-29
ILs	Interleukins
JAK/STAT	Janus Kinase/Signal Transducer and Activator of Transcription
Lcns	Lipocalins

MS	Mass spectrometry
Mda-7	Melanoma differentiation-associated gene-7
T _m	Melting temperature
<i>Mj</i>	<i>Methanocaldococcus jannaschii</i>
MST	Microscale thermophoresis
MAPK	Mitogen-activated protein kinase
mAb	Monoclonal antibody
NAGs	N-Acetyl-d-glucosamines
NK	Natural killer
ncAA	Non-canonical amino acids
NBY	Nitrobenzyl-tyrosine
pSTAT3	Phosphorylated STAT3
PDB	Protein data bank
PKR	Protein kinase A
PPI	Protein-protein interactions
RA	Rheumatoid arthritis
RMSD	Root mean square deviation
scFv	Single-chain variable fragments
SPR	Surface plasmon resonance
SLE	Systemic lupus erythematosus
Th2	T helper 2
3D	Three-dimensional
TB	Tuberculosis
Tyk2	Tyrosine kinase 2
TyrRS	Tyrosyl-tRNA-synthetase
WB	Western blot
WT	Wild-type

List of contents

Declaration.....	
Acknowledgment	
Abstract.....	
Abstract (in Czech)	
List of publications	
List of abbreviations.....	
List of content	
I. INTRODUCTION	1
A. Protein-protein interaction	1
B. Cytokines.....	1
C. Cytokines of interleukin 10 family	2
1. Interleukin-10.....	4
2. Interleukin-24.....	8
3. Interleukin-20 receptor 2.....	12
4. Interferon gamma receptor 2	14
D. Protein scaffold engineering	16
1. Establishment of protein scaffold engineering.....	16
2. Directed evolution based on display technologies	19
3. Ribosome display	21
E. Photoxenoprotein engineering	23
1. Introduction of photoxenoprotein engineering	23
2. Light-responsive non-canonical amino acids.....	23
3. Genetic code expansion technology	25
II. GOALS	28
III. RESULTS	29
A. Develop a general strategy to find new scaffolds and evolve them into IL-10 binder	29
1. Screening of suitable protein scaffold candidates from the PDB	32

2. In silico identification of mutable patches on the surfaces of the scaffold candidates	34
3. Characterization of the solubility and stability of scaffold candidates	35
4. Characterization of alanine mutants, solubility and thermostability	36
5. Final decision on scaffold selection and design of combinatorial library	36
6. Ribosome display with human IL-10	37
7. Characterization of IL-10 binders	38
8. Binder crystallization	39
B. Solving 3D protein structure of IFN- γ R2	40
1. Biophysical characterization of the extracellular domain of human IFN- γ R2	40
2. Structural features of IFN- γ R2	40
3. Structure and sequence alignments revealed important features of cytokines	42
C. Modulating the interaction of IL-24 and IL-20R2 by photoxenoprotein engineering	44
1. Photoxenoprotein design	45
2. Genetic code expansion for photoxenoprotein production	46
3. Purification and characterization of photoxenoproteins	46
4. Decaging efficiency upon UV irradiation	47
5. Monitoring photocontrol by affinity measurement in solution	47
6. Monitoring photocontrol by detection of activated signaling pathway	49
IV. DISCUSSION	51
V. CONCLUSION.....	54
VI. REFERENCES	55
VII. APPENDICES.....	66
Publication 1	66
Supplemented publication	86
Publication 2	92
Publication 3	101

I. INTRODUCTION

A. Protein-protein interactions

Protein-protein interactions (PPI) and their networks are fundamental to many biological processes [1]. Protein homo- and hetero-oligomerization and subsequent formation of biomolecular complexes and organelles without membranes play pivotal roles in our body such as signal transduction, membrane transport, electron transfer, muscle contraction, and others. Aberrant PPI are associated with many human pathological conditions including cancer, infectious diseases, and neurodegenerative diseases [1]. Since drug targets have traditionally been enzymes or ion channels, interference of specific PPIs via small-molecule binding indicated new potential therapeutic targets. However, PPI modulation through small molecules is challenging and PPIs are often regarded as “undruggable” targets [2]. This is due to several reasons. First, the interface area of the interactions is large, usually reaching more than 1500 Å², and highly hydrophobic. Second, the interfaces tend to lack clear grooves or pockets typical for enzyme targets. Third, many amino acid residues are involved in forming the interfaces and the resulting affinities are high. Additionally, intrinsically disordered proteins with highly dynamic structures, which interact with various partners, are also considered “difficult to drug” due to missing binding clefts and the myriad of conformations [3].

B. Cytokines

In the human immune system, communication between cells is critical to regulate the development and growth of the body, to keep self-tolerance, and to protect the body from pathological events such as infectious agents (microorganisms, parasites) and the formation of cancerous cells. A major means of cellular communication is through cytokines.

Cytokines are small, secreted proteins that act as messengers to convert the original intercellular stimulus into intracellular signal. By binding to their cognate membrane-bound receptors on target cell surface, cytokines activate the corresponding signaling pathway, leading to a cascade of changes as a response [4]. Depending on the stimulation, either physiological or pathological, the cellular responses can be different such as changes in gene expression patterns, cytoskeleton reorganization, inflammation, immunomodulatory effects, and release of secretory vesicles [5]. Cytokines are produced by cells of the immune system but also by a variety of non-immune cells including fibroblasts, epithelial cells, endothelial cells and stromal cells [6]. Cytokines can be further classified as interleukins (ILs), interferons (IFNs), chemokines, lymphokines and tumor necrosis factors.

Cytokines are involved in both innate immunity and adaptive immunity of the immune response system. The innate immune system gives the first immediate reaction against any sort of potential invaders. Such a response is very fast but not specific, it has no memory and is not strong. These changes kick-start the innate inflammatory response that establishes a physical barrier to minimize the spread of infection. As the innate immune system is running out of

steam, the adaptive immune system, which is more specific, effective and has memory, takes over. Depending on the type of invader, the immune response is driven to either the humoral immunity or cell-mediated immunity. The first is involved in antibody production of B cells and memory B cells, and the latter is involved in antigen-presenting cell activation and a cytotoxic T cell response.

Cytokine-related disorders lead to several health issues such as autoimmune diseases, allergic diseases, and immunodeficiency diseases. Autoimmune disease occurs when the immune system is hyperactive and attacks the organism's own healthy tissues as if they originated in foreign organisms. Allergy is elicited when the immune response is excessive and harmless antigens are recognized as serious invaders. In the case of immunodeficiency diseases, the immune system is responding insufficiently to infections or malignancies [6].

In the lab, we concentrate on a specific group of cytokines, so called family of interleukin 10. The members of the family are important members of native immunity system. Moreover, these proteins are all potential targets for medical applications. Studying the IL-10 family is our long standing focus and up-to-now the lab research has been published in several papers: IL-10 [7, 8], IL-24 [9, 10], IL-29 [11], IFN- γ [12], [13], [14] [15] and other ILs such as IL-9 [16].

C. Cytokines of the interleukin-10 family

ILs can be classified by various criteria; we will use the system combining functional and structural features of these signaling molecules. The IL-10 family of cytokines is assigned based on the structure homology, shared usage of receptor complexes, and similar and overlapping downstream signaling pathways. ILs belonging to the IL-10 family are produced by both innate and adaptive immune cells. They show indispensable functions as effectors and regulatory molecules of the immune system in cases of infection, inflammation, tissue homeostasis, autoimmunity, and cancer [17, 18].

The IL-10 family includes important immune proteins IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28, and IL-29, which are considered as class two α -helical cytokines. The genes encoding IL-10, IL19, IL-20 and IL-24 map to chromosome 1; IL-22 and IL-26 are located together with IFN- γ on chromosome 12; whereas IL-28A, IL-28B and IL-29 co-localize on chromosome 19. Amino acid sequences of these ILs are not conserved, displaying only 20-30% homology, but their secondary structures are highly similar. They comprise of six α -helices (A–F) and connecting loops, with four helices compacted into a left-handed four-helix bundle [19].

The IL-10 family members modulate their target cells by binding to the extracellular domain of their corresponding membrane-bound receptors in order to activate a cascade of endogenous phosphorylation events, particularly the JAK-STAT signal transduction pathway (**Figure 1**). Receptors consist of two chains (R1 and R2 or RA and RB). Even though all cytokines display high structural homology, shared use of receptor complexes, and have similar roles as important immuno-regulators, they also have unique roles including immune suppression (IL-

10), skin biology (IL-19 and IL-20), tumor suppression (IL-24), antiviral responses (IL-28 and IL-29) [6], and others [20].

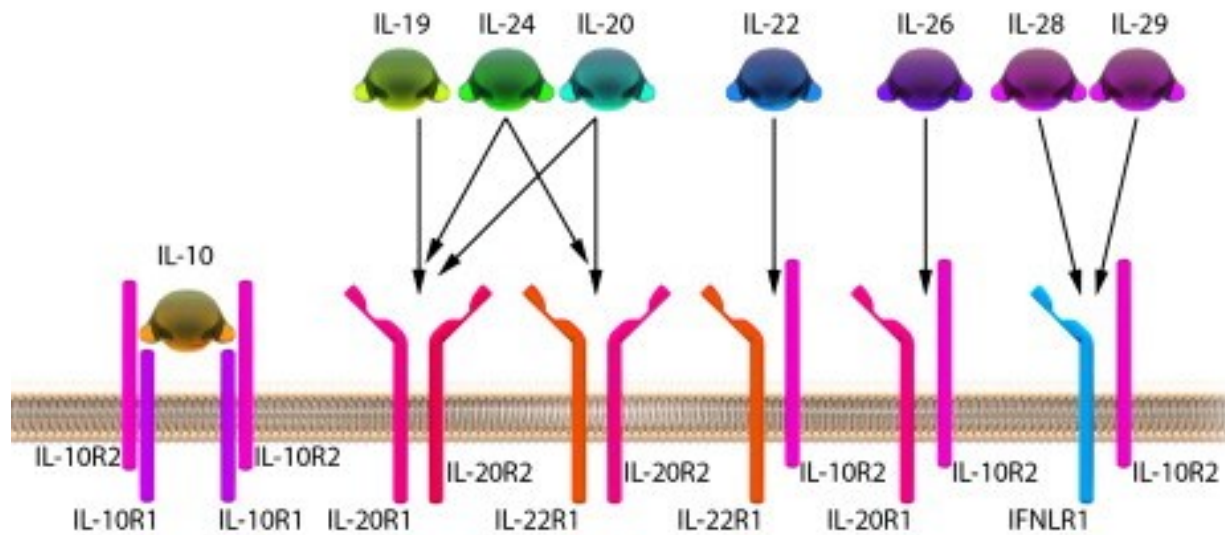


Figure 1: The IL-10 family of cytokines share common receptor subunits. The IL-10 signals via complex of IL-10R1/IL-10R2. The IL-19, IL-24, IL20 signal via complexes of IL-20R1/IL-20R2 and IL-22R1/IL-20R2. The IL-22 signals via complex of IL-22R1/IL-10R2. The IL-26 signals via complex of IL-20R1/IL-10R2. The IL-28 and IL-29 signal via complex of IFN- λ R1/IL-10R2. Reprinted from [6] with permission, copyright © 2011 American Academy of Allergy, Asthma & Immunology. Published by Mosby, Inc. All rights reserved.

1. Interleukin-10

a. Discovery

The anti-inflammatory IL-10, initially described and named as cytokine synthesis inhibitory factor (CSIF) by Fiorentino's group in 1989, was found to be secreted by activated murine Th2 cells. CSIF prevented the production of other ILs and IFN by activated Th1 cells [21]. Later, the human IL-10 was also proved to exhibit CSIF activity [22]. Moreover, both mouse and human IL-10 exhibited extensive homology to an open reading frame from BCRF1 of Epstein-Barr virus (EBV) genome [22, 23], and others such as herpesvirus and poxvirus [24, 25]. These viral homologs are able to bind the human IL-10 receptors and share suppressive properties of human IL-10, suggesting the virus's strategy to suppress host response [26].

b. Gene

The IL-10 gene is mapped on chromosome 1 (q31-32) [27]. It is 2 kb long and contains 5 exons encoding for signal peptide, loops and helix folds [28]. IL-10 gene expression is controlled at the transcriptional level by the transcription factors Sp1 and Sp3 and at the post-transcriptional level through sequences in the 3'-untranslated region [29] [30].

c. Protein structure

The secreted functional unit of IL-10 is a non-covalently linked homodimer, each monomer consists of 178 amino acids with a molecular weight of ~18 kDa. Its X-ray structure is deposited in the protein data bank (PDB) as ID 2ILK (**Figure 2**) [31].

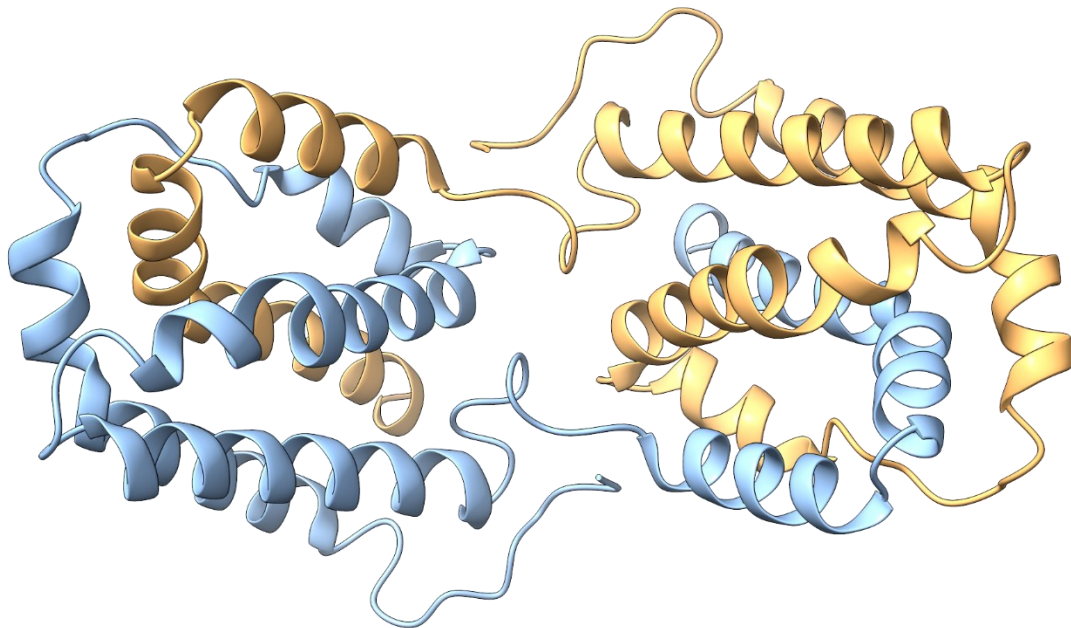


Figure 2: Structure of human IL-10 dimer. The two monomers (in ribbon representation) are distinguished in blue and yellow colors. Figure created with ChimeraX from PDB ID 2ILK [31].

d. Receptor interaction

The IL-10 communicates with target cells by binding to its specific receptors IL-10R1 and IL-10R2 on the cell surface. The IL-10 homodimer binds to a hetero-tetrameric receptor complex made up of two IL-10R1 and two IL-10R2 chains [32], with high affinity (dissociation constant of 50-200 pM) [33]. The crystal structure of the complex between IL-10 and IL-10R1 reveals that two IL-10 molecules bind to the IL-10R1 dimer. A monomer of IL-10 contains 6 α -helices: 4 helices are donated from one chain and 2 helices are donated from another asymmetric chain. A subunit of soluble IL-10R1 contains two domains and each one has a β -barrel like structure. The structure is deposited as 1J7V on the PDB [34]. The structure of the complex between the IL-10 and its receptors has not been determined yet.

e. Signal transduction

The intracellular part of IL-10R1 and IL-10R2 are bound to Jak1 and Tyk2 constitutively [35, 36]. Upon binding of IL-10 to the receptor complex, Jak1 and Tyk2 are activated [37], followed by phosphorylation of STAT1, STAT3, and STAT5 [38]. As a result, SOCS1 and SOCS2, known as signaling inhibitors, are expressed [39].

f. Functions

- The IL-10 is an important immune regulator due to its immunosuppressive, anti-inflammatory and stimulatory effects in immune system.

The IL-10 is a pleiotropic cytokine that is produced by various myeloid- and lymphoid-derived immune cells participating in both innate and adaptive immunity. Mainly, IL-10 exerts its immunosuppressive functions to protect tissue damage caused by exaggerated and uncontrolled inflammatory effector responses, especially during the resolution phase of infection and inflammation. It also functions to maintain homeostasis with gut microbes [18].

As an immune suppressor, the IL-10 downregulates the function of activated monocytes/macrophages and dendritic cells via repressing their production of pro-inflammatory cytokines and antigen-presenting capacity. Moreover, by diminishing the antigen-presenting capacity of monocytes via downregulation of MHC class II surface expression, IL-10 indirectly prevents excessive antigen-specific human T cell activation and proliferation [40-44].

The anti-inflammatory effects of IL-10 are driven by binding to its cognate receptors highly expressed on monocytes/macrophages. The activated Jak1-Tyk2-Stat3 cascade results in STAT3-mediated transcription of genes encoding proteins such as Bcl-3 and Ets-3 that suppress the expression of pro-inflammatory genes. As a result, the anti-inflammatory response is limited [45].

In contrast to its inhibitory effect, the IL-10 also has a stimulatory effect in B cells. IL-10 enhances the survival of human B cells as well as their proliferation, differentiation into plasma cells, and isotype switching [46].

- IL-10 plays various roles in autoimmunity, allergy, cancer and infection.

Since IL-10 is a multifunctional cytokine being both a potent immune-suppressor and immune-stimulator, its expression must be tightly regulated. Otherwise, altered IL-10 expression levels, identified by SNPs and microsatellites, can lead to the development of a large number of serious diseases including autoimmune disorders, cancer, allergies and infectious diseases [6].

- *IL-10 deficiency results in unrelenting immune activation [47].*

This is the case in chronic inflammatory bowel diseases [48] (e.g. Crohn's disease [49]), psoriasis [50], rheumatoid arthritis [51], and after organ transplantation [52]. Given its multiple anti-inflammatory functions as well as its repression of pro-inflammatory cytokines, IL-10 is a potential therapeutic target for the treatment of autoimmune and inflammatory diseases. Clinical studies showed that administration of recombinant IL-10 provided benefits in patients with Crohn's disease [53], psoriasis [50], and rheumatoid arthritis [54].

The IL-10 may also represent a potential anti-tumor immunotherapy in human patients with cancer because IL-10 stimulates the expansion and cytotoxicity of tumor-infiltrating CD8+ T cells, which play a key role in killing tumor cells. Mice deficient in IL-10 show weakened tumor immune surveillance but recovering IL-10 levels by overexpression protected transgenic mice against cancer development. Clinical trials of cancer patients treated with IL-10 showed that IL-10 activates tumor-infiltrating CD8+ T cells, therefore limiting the development of solid tumors [55-57] (**table 1**).

The IL-10 has a protective role in allergic diseases. Experimental evidence suggests that IL-10 expressed by APC cells in the respiratory tract is reduced in patients with asthma. Moreover, allergen-specific T helper 2 (Th2) cells and their cytokines promote the allergic airway inflammation [58]. Therefore, one of the strategies for the treatment of allergic patients is to use IL-10 producing dendritic cells to downregulate allergen-specific T helper 2 responses [59].

- *The IL-10 over-production results in unwanted immunosuppressive effects, tumor growth and infection [47].*

The IL-10 has a pathogenic role in autoimmune diseases, like systemic lupus erythematosus (SLE), since IL-10 was reported also to be a growth factor for human B cells, to promote antibody production, class switching, and plasma cell differentiation [60, 61]. The severity of SLE correlates with high amounts of serum IL-10 and autoantibodies [62]. Using anti IL-10 mAb to treat patients with SLE was proved effectively in clinical trials [63] (**table 1**). These studies showed that developing IL-10 antagonists helps to control SLE pathogenesis and therapy.

Epstein-Barr virus latently infected B-cells are the precursors of EBV-associated malignancies. EBV not only uses its IL-10 homolog that binds IL-10 receptor to increase proliferation [64] and protect infected EBV-infected B-cells from immune recognition [65], but also enhances host IL-10 production to promote cell survival and development of associated malignancies [66].

In infectious diseases, elevated IL-10 production have been found in lungs and sera of patients with advanced tuberculosis [67]. This suggests a potential use of IL-10 antagonists for hard-to-treat patients with tuberculosis.

A paradoxical dual role of IL-10 in cancer has been presented with both tumor-repressing and tumor-promoting functions [56, 68].

Cytokine	Intervention	Indication	Clinical stage	Sponsor
IL-10	Tenovil (rhIL-10)	Crohn's disease	Phase I/II N/A	Schering-Plough Research Institute
	Dekavil (F8-IL-10)	Rheumatoid arthritis	Phase II NCT02076659 Phase II NCT02270632	Philogen/Pfizer
	Tenovil TM (IL-10)	Acute pancreatitis	Phase II NCT00040131 (terminated)	Merck Sharp & Dohme
	IL-10	Psoriasis	Phase II NCT00001797	National Cancer Institute
	Prevascar (rhIL-10)	Cicatrix, wound healing	Phase II NCT00984646	Renovo
	AG011 (engineered Lactococcus lactis secreting human IL-10)	Ulcerative colitis	Phase I/II NCT00729872	ActoGenix N.V.
IL-10	AM0010: PEGylated human IL-10	Solid tumors/pancreatic cancer	Phase I NCT02009449 Phase III NCT02923921	ARMO BioSciences
	BT063 (antibody to neutralize IL-10)	Systemic lupus erythematosus	Phase II NCT02554019	Biotest
	INGN 241 (Ad-mda-7)	Melanoma	Phase II NCT00116363	Introgen Therapeutics

Table 1: Clinical trials targeting interleukin IL-10 and IL-24. Adapted from [69] with permission.

2. Interleukin-24

a. Discovery

The IL-24 gene was initially named melanoma differentiation-associated gene-7 (mda-7) by Paul Fisher's group in 1995. In their experiments, fibroblast interferon and mezerein were used to treat melanoma cells to terminate differentiation and suppress growth and tumorigenic abilities. Through subtractive hybridization of cDNA libraries, the mda-7 mRNA levels were found to be elevated in melanoma cells compared to non-induced cells. Observing that the treatment of human melanoma cells results in growth suppression and increased mda-7 expression, suggested that the mda-7 gene may function as a negative regulator of melanoma progression [70]. Moreover, mda-7 was proposed as a novel tumor suppressor in various cancers [71, 72].

The Mda-7 gene is mapped on human chromosome 1q32. Human mda-7 consists of seven exons and six introns encoding a protein with 206-residues, 49 of them serve as an N-terminal signal peptide for protein secretion. Primary structure analysis also shows IL-10 sequence signature from amino acids 101-121 and phosphorylation sites. Moreover, there are three putative N-glycosylation sites at amino acids 95, 109, and 126 resulting in the expression of several isoforms with molecular weights ranging from 18.3 kDa to 35 kDa. Besides, protein expression profiling studies showed that Mda-7 is present in many immune cell types such as thymus, spleen, and peripheral blood leukocytes [73, 74].

Through sequence homology analysis, rat and mouse orthologues of human mda-7 were identified. The rat C49a gene was found to be associated with fibroblast proliferation in wound healing [75] and the rat Mob5 gene was shown to encode a cytokine-like secreted protein activated by the Ras oncogenes [76]. FISP protein was found as a secreted IL-4-induced-protein from Th2 cells [77].

Taken into account its chromosomal location, expression profile and cytokine-like properties, mda-7 was eventually renamed to IL-24 by the Human Gene Organization (HUGO) [78].

b. Protein structure

The structure of human IL-24 in a ternary complex with the extracellular domains of two receptors (IL-22R1 and IL-20R2) has been solved by X-ray crystallography at 2.5 Å resolution and deposited on PDB as entry 6DF3 [79]. The complex was formed between the high affinity receptor IL-20R2 and IL-24 fused to the low affinity receptor IL-22R1 through a peptide linker. IL-24 has a 4-helix bundle fold even though it contains 6 helices; without any disulfide bond [79]. Later, the structure of a highly mutated receptor-free human interleukin IL-24, which was solved at 1.3 Å resolution and deposited to PDB as 6GG1 by our lab [9]. The structure of this isolated engineered IL-24 showed similarity to the native IL-24 in the complex form (**Figure 3**), except for the formation of a disulfide bond. As noted, native IL-24 and engineered IL-24 were produced in different systems, *Drosophila* S2 cells and *E. coli*, respectively [9].

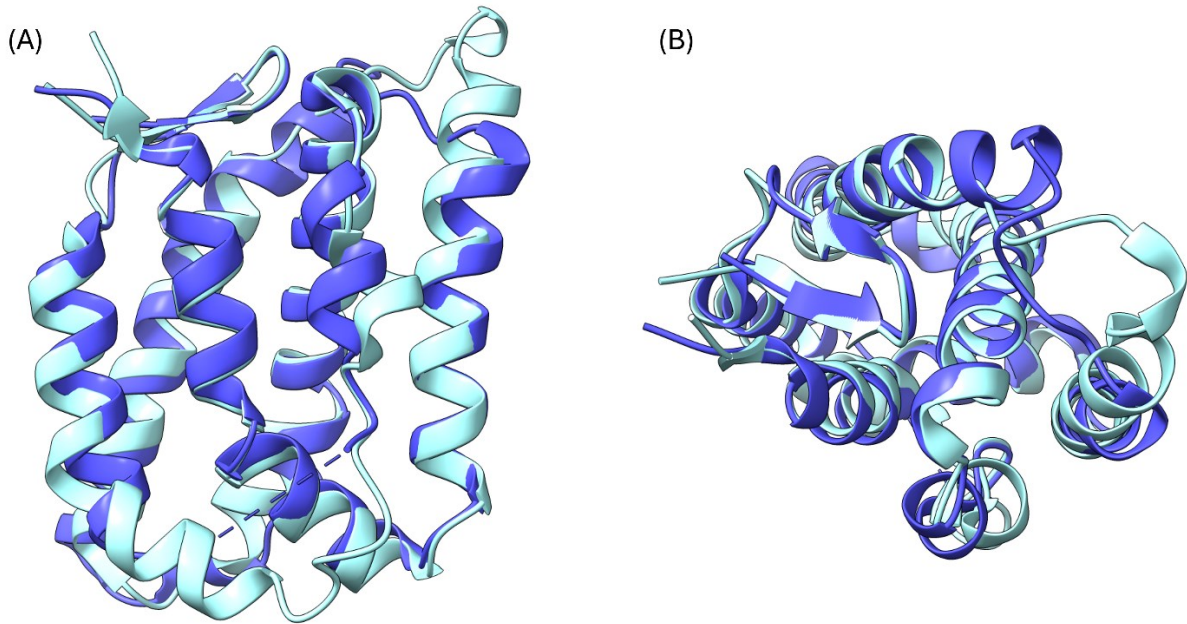


Figure 3. Superimposition of two IL-24 structures: the engineered IL-24 (PDB ID 6GG1) [9] in blue over the native IL-24 (PDB ID 6DF3) [79] in cyan. (A) Side view. (B) Top view. Figure created with ChimeraX from these two PDB IDs.

c. Receptor interaction

There are three membrane-bound receptors for IL-24 comprising IL-22R1, IL-20R1, and IL-20R2. IL-24 signals through heterodimers made of combination of these receptors IL-20R1/IL-20R2 and IL-22R1/IL-20R2. IL-20R2 is present in both receptor complexes. While IL-20R1/IL-20R2 dimer is shared among IL-19, IL-20, and IL-24; the IL-22R1/IL-20R2 dimer is shared between IL-20 and IL-24. Therefore, the IL-24 belongs to IL-20 subfamily of cytokines [74, 80, 81].

d. Signal transduction

Like all IL-10 family cytokines, the IL-24 triggers signaling via the Tyk2/Jak1 pathway, which continuously activates downstream transcription factors (STAT1/STAT3) through phosphorylation [82]. This pathway is regulated by SOCS proteins [83]. Upon binding of IL-24 to the extracellular domains of its heterodimeric receptors and forming a hetero-trimer complex, the IL-24 signals via JAK-STAT pathway in epithelial cells [74] and in ovarian carcinoma cell line [81].

In addition to this well-established JAK/STAT-dependent mechanism, different JAK/STAT-independent mechanisms have been reported [84]. For instances, IL-24 induces apoptosis in melanoma, breast, fibrosarcoma, and prostate cancer cell lines independently of the STAT3 pathway [85, 86]. Moreover, non-secreted intracellular IL-24 protein is also able to induce apoptosis in prostate cancer cell lines [87]. In these cases, IL-24 binds to distinct intracellular molecular partners such as protein kinase A (PKR), sigma 1 receptor, and respiratory chain protein Grim19 to activate PKR-driven inflammatory response [88], cancer-specific apoptosis in the endoplasmic reticulum and plasma membrane [89], and recruitment of STAT3 to

mitochondria [90], respectively. Additionally, signaling of IL-24 through other JAK/STAT-independent mechanisms have been discovered, such as the p38 mitogen-activated protein kinase (MAPK) pathway [84]. These results indicate the complexity of IL-24 signaling in the regulation of cellular responses (**figure 4**) and the potential importance of IL-24 as a drug target [84].

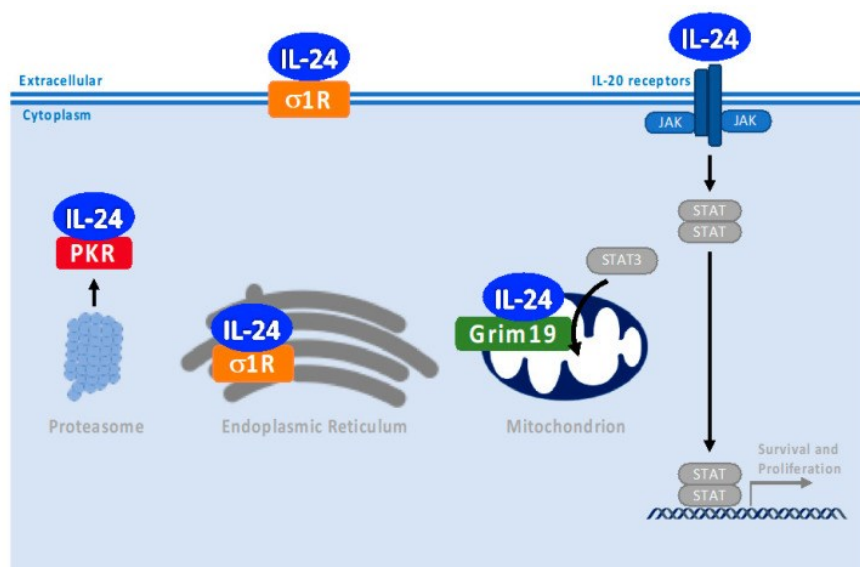


Figure 4. The interaction partners of IL-24: the IL-20 receptors at the plasma membrane, which trigger the well-established JAK/STAT-dependent signaling mechanism, and several proteins found in different cellular compartments, which give rise to alternative JAK/STAT-independent mechanisms. Reprinted from [84] with permission.

e. Functions

Nowadays, the IL-24 is considered an effective cancer-killing cytokine due to its potent anticancer effect, and a pleiotropic immuno-modulatory cytokine due to its relevance in immune response, host defense, and tissue homeostasis (reviewed in [91, 92]).

- The IL-24 is a target in cancer therapy.

The antitumor role of IL-24 has been extensively studied and confirmed in diverse human cancer cells, such as inhibition of cancer invasion, metastasis, angiogenesis and cancer stem cell growth; and induction of cancer apoptosis, and autophagy (**figure 5**) [93-98]. IL-24 became a therapeutic target in Phase I/II clinical studies using intra-tumoral injections of an inactivated adenovirus containing the mda-7 construct [Ad-mda7; INGN-241] [99].

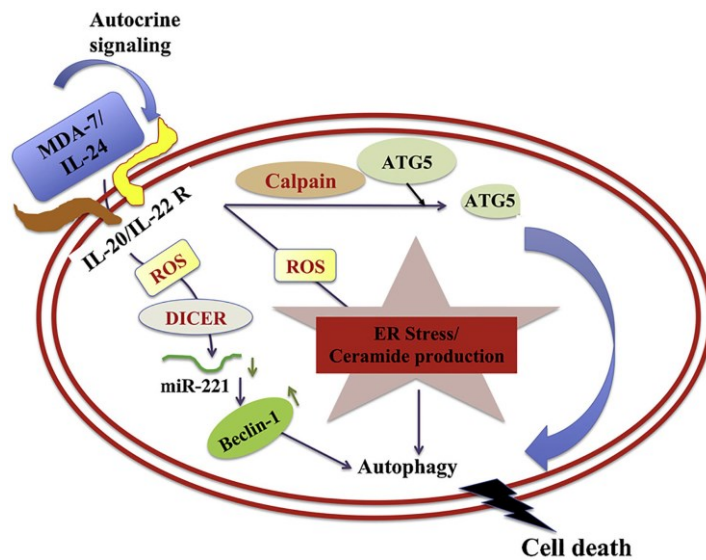


Figure 5: IL-24 regulates autophagy and subsequent cell death through interaction with its receptors. Reprinted from [100] with permission, © 2019 Published by Elsevier Ltd.

Meanwhile, the unique aspects of IL-24 antitumor “bystander” activity has been observed in many studies from cell, preclinical animal models and in a Phase I/II clinical trial in patients with advanced cancers. Transfection of normal cells with IL-24 gene creates a secreted IL-24 protein product that can effectively activate apoptosis on distant tumors, which did not initially receive this gene product, through IL-24 dimeric receptor pairs on the cell surface [86, 101].

Interestingly, it has also been reported that IL-24 fragments can be separately transcribed as different splice variants which cause tumor apoptosis by inducing several different pathways [102].

Several types of gene therapy methods relying on transfection of cells with IL-24 gene, and of protein-based therapy method aimed for sufficient penetration of recombinant IL-24 protein into tumor have been investigated and developed. They are virus-mediated gene delivery [99], T cells expressing IL-24 [103], recombinant IL-24 protein [98, 104] and nanoparticle-mediated delivery [105].

- Therapeutic Potential of Targeting IL-24 in immunity.

The IL-24 is mainly produced and secreted by both immune and non-immune cells such as T cells, B cells, monocytes, macrophages, NK cells, keratinocytes, and fibroblasts [106]. The IL-24 displays various functions in autoimmunity, inflammation, and infectious diseases.

- Depending on the type of autoimmunity and the site of inflammation, IL-24 exhibits both pro-inflammatory or anti-inflammatory properties.

The IL-24 displays a pathological role in autoimmune diseases such as psoriasis or rheumatoid arthritis (RA). Psoriasis is a chronic inflammatory disease of the skin resulting from the excessive

proliferation and differentiation of keratinocytes. Elevated IL-24 and its IL-20R2 receptor levels plus its stimulatory effect of pro-inflammatory molecules causing skin lesions, suggest that IL-24 might be a factor in the pathogenesis of psoriasis, particularly promoting skin inflammation [107]. RA is an inflammatory autoimmune disease that can lead to progressive joint damage and disability. Increased IL-24 levels are observed in RA patients and have been proved to promote mononuclear cell infiltration, which is a hallmark of chronic inflammation. Moreover, anti-inflammatory treatment of RA patients decreased IL-24 levels [108].

However, the IL-24 can also play a protective role in autoimmune diseases. Inflammatory bowel disease (IBD) is a chronic inflammation of the gastrointestinal tract. Studies showed increased expression of IL-24 in inflamed mucosa of active IBD patients. IL-24 targets colonic epithelial cells to enhance JAK/STAT1 phosphorylation and expression of SOCS3 and mucins, supporting their suppressive effects of mucosal inflammation in IBD [109].

- *The IL-24 is a potential treatment for infectious diseases.*

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis*. Significantly, lower levels of MDA-7/IL-24 have been found in patients with active TB compared to those with a latent TB infection. Since IL-24 was shown to regulate IFN- γ in TB patients and IFN- γ is well known for therapy of TB patients, IL-24 might be a treatment option for these patients [110].

3. Interleukin-20 receptor 2

The IL-20R2 is a transmembrane protein detecting IL-19, IL-20, and IL-24 by its extracellular region and transducing the signal across the membrane to intracellular JAK phosphorylases. IL-20R2 pairs with either IL-20R1 or IL-22R1 to form heterodimeric cytokine receptors IL-20R1/IL-20R2 or IL-22R1/IL-20R2. The IL-20R1/IL-20R2 receptor complex is shared among IL-19, IL-20, and IL-24, while the IL-22R1/IL-20R2 receptor complex is a shared between IL-20 and IL-24 [18]. Therefore, IL-20R2 is the common subunit of both receptor complexes. Its 3D structure from PDB ID 6DF3 is illustrated in **Figure 6** [79].

Binding between the extracellular domain of the receptors, either IL-22R1/IL-20R2 or IL-20R1/IL-20R2, and their corresponding cytokines [74] induces signaling cascade in the target cells through JAK/STAT pathway, which subsequently activates the expression of downstream transcription factors particularly STAT1 [85], STAT3 [111] and STAT5 [112].

The IL-20 receptors, which are expressed by various cell types, play a role in the reciprocal interaction between immune and non-immune cells including keratocytes, epithelial and endothelial cells [20]. The communication between leukocytes and epithelial cells is facilitated, thus enhancing defense and tissue repair processes [113].

Even though IL-19, IL-20, and IL-24 display structural homology, signal through the same set of receptors, and induce the same signaling pathways, they also display unique biological functions in immune regulation, tissue homeostasis, host defense, and oncogenesis. This could

be explained by the different binding affinities to their shared receptor complexes, which could induce distinct signaling and regulation [20]. Despite the promiscuity in the receptor use of IL-19, IL-20, and IL-24 (**Figure 1**), their physiological effects appear quite distinct: immune activity with IL-19, skin biology with IL-20, and tumor apoptosis with IL-24 [111].

The elevated expression of IL-19, IL-20 and IL-24 in patient samples are associated with immunity-related diseases such as psoriasis, RA, and IBD diseases. In contrast to their pathogenic role in skin inflammation found in psoriasis and RA, they play a protective and even therapeutic role in inflamed intestine in IBD [113]. Besides, the three cytokines are also involved in promoting *Staphylococcus aureus* infection in mice and humans. *S. aureus* gives rise to the majority of human skin and soft tissue infections, and is a major infectious cause of mortality. Neutralization of IL-20R2 by antibody improves the outcome of bacterial infections and changes the susceptibility to infection [114].

The IL-20R2 receptor expression levels have been reported in muscle invasive bladder cancer [115], and papillary renal cell carcinoma [116]. A study revealed that IL-19 stimulated IL-20R2–expressing tumor cells to raise tumor cell proliferation in bone and bone metastasis. Therefore, a new strategy to cure lung cancer bone metastasis could be based on IL-20R2 blockage [117].



Figure 6: Structure of IL-20R2. Figure created with ChimeraX from PDB ID 6DF3 [79].

4. Interferon gamma receptor 2

The IFN- γ R2 is the crucial cell-surface receptor of IFN- γ . A matured 310 amino acid IFN- γ R2 comprises three domains including an intracellular domain, a transmembrane domain and an extracellular domain. This extracellular domain consists of five cysteines and six potential N-linked glycosylation sites making its molecular weight in the range of 61 to 67 kDa [118].

The IFN- γ R2 plays an important role in IFN- γ signaling, influencing innate and adaptive immunity against pathogens and tumors [119, 120]. Based on the analysis of its sequence similarity, it is classified in the class two receptor family including IFN- γ R1, receptors of interferon- α and interferon- β (IFN- α and IFN- β), receptors of IL-10 and IL-20, and receptors of other interleukins belonging to the IL-10 family [121]. The IFN- γ homodimer signals through forming a ternary complex with receptors IFN- γ R1 and IFN- γ R2 subunits, consequently activating the JAK/STAT signaling cascade. The IFN- γ binds to its high affinity receptor IFN- γ R1 to form a binary complex. Subsequently, the IFN- γ R2 must interact with this binary complex to form a ternary complex [122-124]. The signaling pathway is shown in **Figure 7A**.

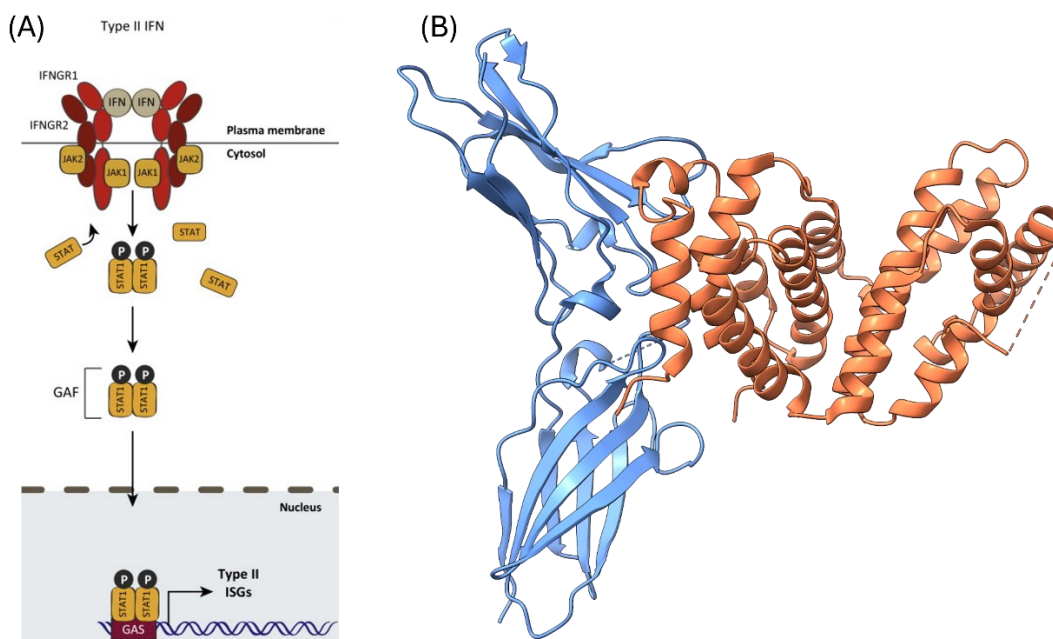


Figure 7:

(A) IFN- γ signals via the JAK/STAT pathway upon binding to its two membrane-bound receptors (IFN- γ R1 and IFN- γ R2) and formation of a ternary complex. Reprinted from [125] with permission, copyright © 2015 Elsevier Ltd. All rights reserved.

(B) The structure of the binary complex consists of an IFN- γ homodimer (orange) and IFN- γ R1 (blue). Figure created with ChimeraX from PDB ID 1FYH [126].

Before my paper had been published, the structure of a binary complex (**Figure 7B**) comprising a homodimer IFN- γ bound to its high affinity IFN- γ R1 had been solved, PDB ID 1FYH [126] and

PDB ID 1FG9 [127], but the detailed biophysical and structural characterization of IFN- γ R2 was missing.

After my paper where we solved the structure of IFN- γ R2 (PDB ID 5EH1) [14], the structure of the ternary complex comprising two IFN- γ , two IFN- γ R1 and two IFN- γ R2 was solved and deposited as PDB ID 6E3K [128] (**Figure 8**).

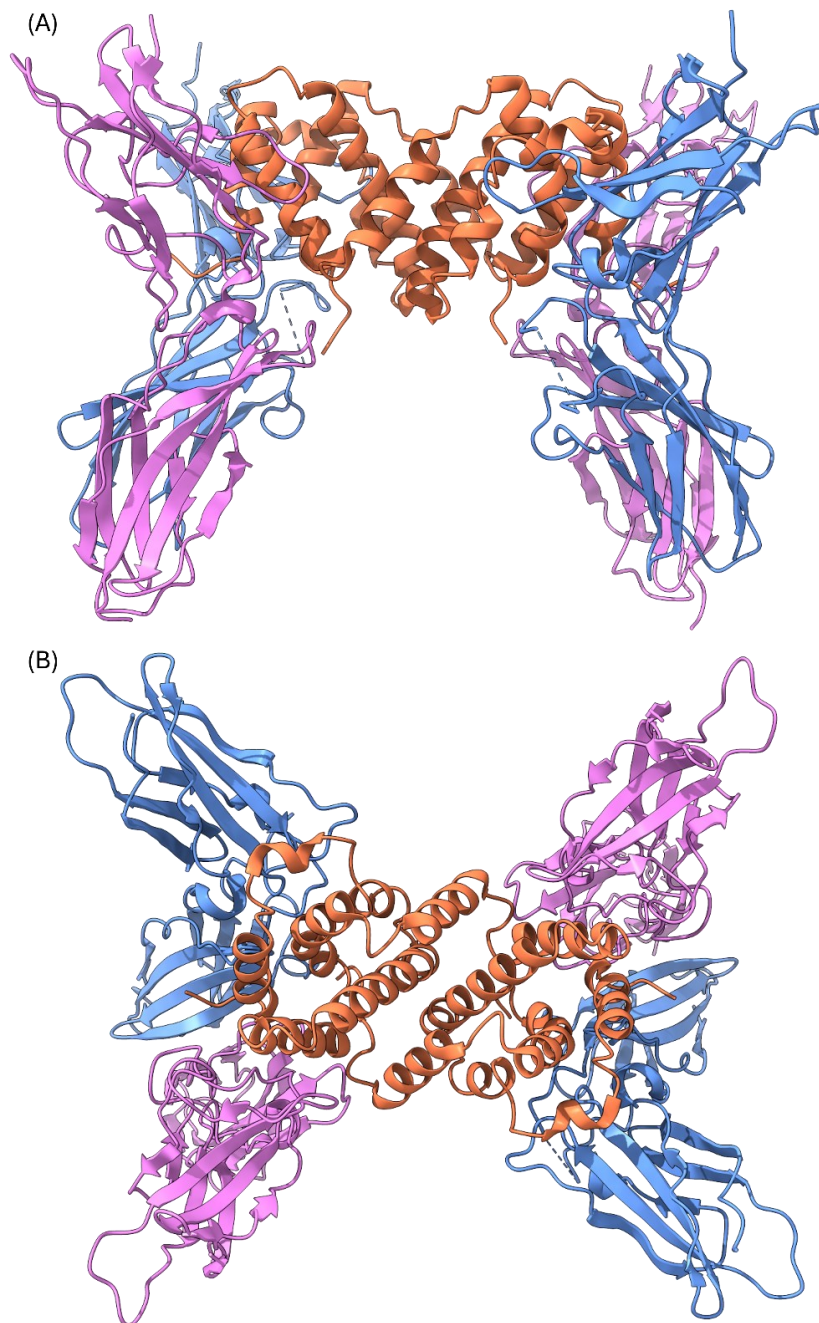


Figure 8: The ternary complex of IFN- γ consists of two IFN- γ molecules (orange), two IFN- γ R1 (blue) and two IFN- γ R2 (pink). (A) Side view of the complex. (B) Top view of the complex. Figure created with ChimeraX from PDB ID 6E3K [128].

D. Protein scaffold engineering

1. Establishment of protein scaffold engineering

a. Antibody engineering

- Monoclonal antibodies (mAb)

Antibodies help us to effectively fight pathogens because of their capacity to neutralize antigen function via direct interaction. In 1975, hybridomas, which are hybrid cells arising from the fusion of mouse antibody-producing B-lymphocytes with mouse myeloma from an immunized donor, were established to produce highly specific monoclonal antibodies in vitro in massive cultures for the first time [129]. The hybridoma technology became a gateway for generating tens of thousands of research tools and therapeutic antibodies [130]. However, it has been reported that the DNA/protein sequence information of monoclonal antibodies can get lost in case hybridoma cell lines die, and different production batches lead to distinct antibody specificity. As a result, experiments using antibodies generated by hybridoma technology may not always be reproducible [131].

- Recombinant antibodies

The rapid growth of recombinant DNA technology, x-ray crystallography, and computer modeling of protein structure and folding, has had a positive impact on the field of protein engineering [132]. Protein production and engineering can also be applied to antibodies. The first solved antigen-binding fragment (Fab) crystal structures in 1970s illustrate the variable and constant parts of immunoglobulins and reveal the “Ig-fold” [133, 134]. Moreover, the positions of hypervariable amino acid sequences are displayed and are believed to furnish the conformation and antigen-binding specificity [133, 134]. Simultaneously, the recombinant DNA technology, which also started in early 1970s, was first applied to the production of recombinant antibodies, a chimeric mouse-human Ig antibody, in 1984. This chimeric antibody consists of Fab of immunized-mouse monoclonal antibody and the constant fragment (Fc) region genes of human immunoglobulin [135]. In contrast to monoclonal antibody production by hybridoma technology, using recombinant DNA techniques these two fragments are inserted into a plasmid and expressed in the target host cell such as bacteria or eukaryotic cell lines [135].

- Antibody derivatives

In 1980s, phage display represented a major technological breakthrough due to its ability of physical coupling between the desired phenotype (e.g. high affinity and high selectivity binding protein) and the genotype (DNA sequence) [136]. In 1990, the single-chain variable fragments (scFv), derived in immunized mice against hen egg-white lysozyme, were successfully displayed on filamentous phages and showed to bind to antigens [137]. This achievement opened prospects of antibody engineering. The application of antibody engineering along with

combinatory library techniques resulted in the successful selection of novel high-affinity binders bypassing the need for antigen immunization.

In contrast to the advantages offered by mAbs, the properties of antibody derivatives still have several limitations. The large size (approximately 150 kDa molecular weight) limits their efficient tumor and tissue penetration [138] and hinders the binding of intracellular antigenic epitopes [139]. Furthermore, post-translational modifications, such as disulfide bonds and/or glycosylation, require costly and time-consuming eukaryote cell cultures. In addition, it may be difficult to add chemical handles and linkers for further drug conjugation in an efficient and site-specific manner [140]. Moreover, the mix of heavy and light chains demands complicated genetic constructions. Last but not least, the missing Fc region is important to prolong antibody circulation in blood because of Fc binding to various cell types [141].

Therefore, the motivation to overcome such limitations resulted in antibody engineering targeting shortened antibody variants, including the variable domains of heavy and light chain in Fab region. Since these domains include the antigen binding segments, their antigen neutralization properties are similar to full-length antibodies. Ig derivatives such as scFv (25 kDa), and single domain antibodies (VHH/nanobody) (15 kDa) satisfying the need of both small size and absence of disulfide bond formation are nowadays commonly used in antibody engineering. Their small size facilitates tumor penetration [142]. Moreover, combinatorial antibody technologies provide a rational strategy to isolate antibodies with tailored properties against virtually any target, both in basic research and for medical therapy [143]. Engineered domains keep a typical β -barrel fold (Ig-fold). ScFv antibodies show the most promising traits exploitable for application in human medicine and basic research [143].

Disadvantages of small-size Ig derivatives include their short half-life character in vivo, and poor retention times in the target tissues [142, 144]. Besides, unexpected effects of the truncated scFv antibodies and nanobodies of non-human origin may occur, such as lower avidity due to reduced binding sites and the required humanization processing for reducing immunogenicity, respectively [145].

b. Protein scaffold engineering

- The idea of protein scaffold

The concept of engineered protein scaffolds for molecular recognition arose in the late 1990s after the first successful selection of antibodies after phage display [146, 147]. The concept of protein scaffolds is simple: scaffolds are structurally stable (robust) polypeptide frameworks that need not be based on antibodies. Indeed, several detrimental aspects of both traditional mAbs and engineered antibodies prompted the development of non-Ig binding proteins [142, 144, 145].

Protein scaffolds are binding proteins whose architecture is not necessarily an Ig fold, and their primary structure can be mutated (evolved) to render useful high affinity protein binders for

various biotechnological applications. Protein scaffolds are expected to keep the advantageous features of antibodies, i.e., constant regions that hold/maintain the overall folding stability and variable regions that provide plasticity responsible for high affinity and specificity binding. However, in contrast to mAbs, protein scaffolds are usually small molecules (less than 200 residues), lack disulfide bonds, have a stable fold, and are easy to produce, ideally in bacterial expression systems [141, 147].

Based on the 3D fold, there are two categories of non-antibody protein scaffolds. Those that resemble the β -sandwich structure of Ig domains, and those that have structures different from the Ig domains. In the case of Ig-fold protein scaffolds, they contain loops analog to the complementarity-determining regions (CDRs) of Igs. Such scaffolds, e.g. monobodies and anticalins, are targeted for randomization to create antibody-like binding libraries. On the other hand, there are non-Ig-fold protein scaffolds, which differ from the Ig-fold, such as DARPins, affibodies, Kunitz scaffold, etc (**Figure 9**).

There are two major strategies to create binding sites on the scaffolds. The first one is to modify an existing binding site and the second is to implement a new binding site for the target. During the last 15 years, more than 50 different protein scaffolds have been developed [148]. Here is a summary of currently used non-antibody protein scaffolds:

- *Ig-fold protein scaffolds*

These are protein scaffolds that closely resemble the β -sandwich structure of Ig domains. Monobody/Adnectin scaffold is based on the tenth extracellular domain of human fibronectin III (10Fn3) whose natural target is integrin. 10Fn3 adopts an Ig-like β -sandwich fold with loops at one end but lacks the central disulfide bridge. Such loops, which are similar to the third CDRs of an Ig variable domain, are targeted for randomization in scaffold design [149]. Another example is the anticalin scaffold, which is based on the lipocalins (Lcns). Lcns naturally bind small ligands at a ligand pocket surrounded by four structurally variable loops at the open end of an 8-stranded β -barrel. In scaffold design, this binding interface of Lcns is randomized to generate antibody-like variants [150].

- *Non-Ig-fold protein scaffolds*

These protein scaffolds have various folds and are of different sizes. The well-established DARPin scaffold is based on a consensus sequence of the abundant ankyrin repeat proteins (ARPs). The ankyrin repeat fold provides amino acid positions in structurally rigid regions that can be randomized to make binding surfaces for a repertoire of proteins. Another example is the affibody scaffold based on the Z domain of the staphylococcal surface protein A, which naturally binds the Fc of Igs. The affibody structure is a three α -helix bundle. Two α -helices of the bundle are diversified while generating the library [151].

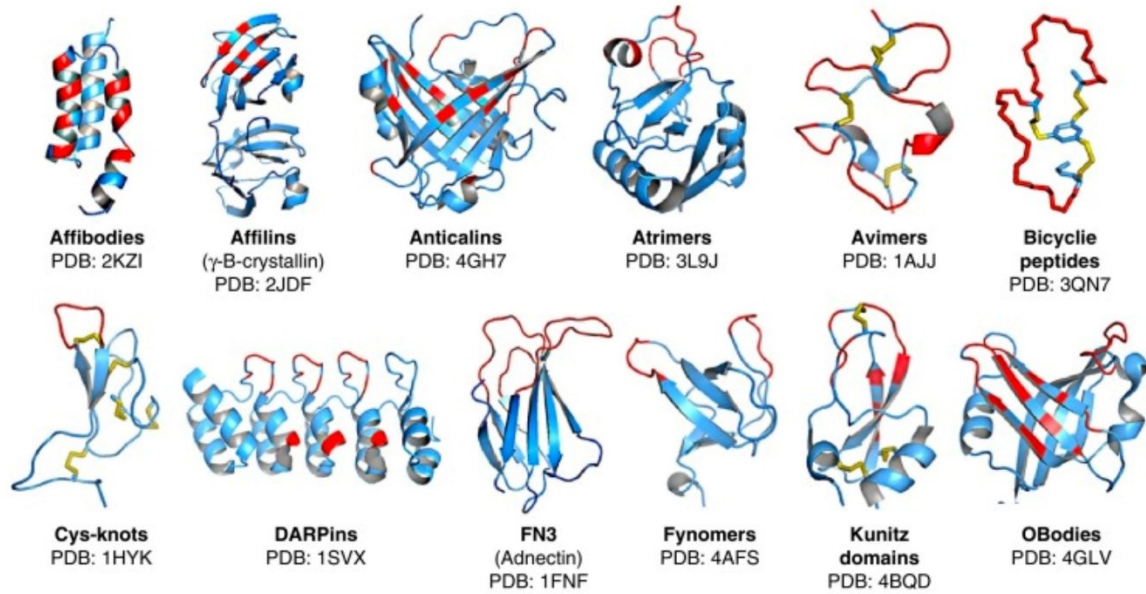


Figure 9: Structures of non-antibody protein scaffolds. Positions targeted for randomization are highlighted in red, and disulfides in yellow. Source [152].

- Antibody and protein scaffold achievements

Until 2015, six products of recombinant antibody therapeutics were granted first marketing approvals and 39 mAbs candidates were in phase III clinical trials as therapeutics [153].

Many small scaffolds have also proven themselves useful in a broad range of applications, from laboratory research tools to imaging, diagnostics, and therapeutics [154-156]. Some of these scaffolds have entered clinical trials [152, 157].

2. Directed evolution based on display technologies

Directed evolution of proteins is a variant of engineering techniques that mimics the natural evolution of protein sequences. Natural evolution of proteins have existed since the emergence of life on Earth. Genes are mutated and their corresponding proteins evolve to adapt to environmental challenges. Directed evolution has been newly established in laboratory settings to speed up the evolution of enzymes and binding proteins for basic research (e.g. aminoacyl tRNA-synthetases for genetic code engineering) and human therapeutics (e.g. design of antibodies and protein scaffolds) [136].

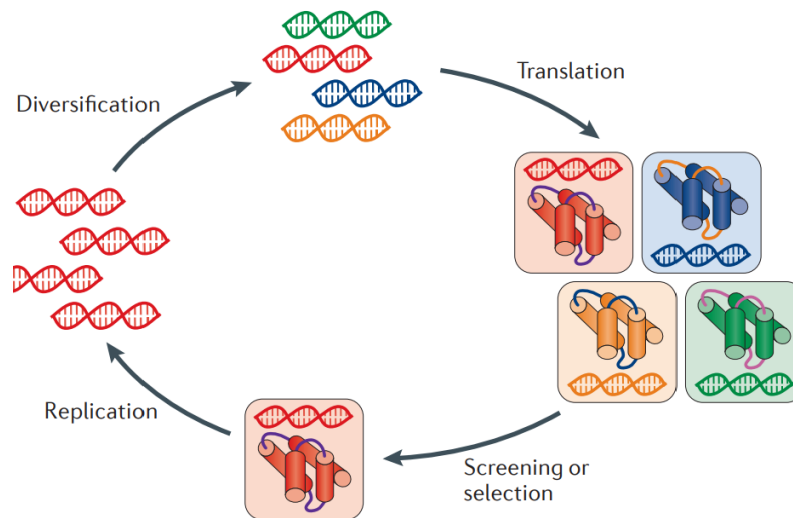


Figure 10: The four key steps in a cycle of directed evolution. Reprinted from [158] with permission, copyright © 2015, Springer Nature Limited.

Directed evolution of binding proteins is a powerful method to identify high affinity and high specificity binders for a given target. For this purpose, the selection of protein binders is facilitated by display technologies that provide a physical link between phenotype (high affinity and high selectivity binding protein) and genotype (DNA sequence). There are two basic types of display technologies: cell-based display (such as yeast display, bacteria display and phage display) and cell-free display including ribosome display and mRNA display. In cell-based display, the genetic information (DNA) of a binder is packed inside a cell or virus (yeast, bacteria or phage), and the translated binders are displayed on the surface for binding assays against target proteins. In cell-free display, the gene information (mRNA) of a binder is associated physically to a translated binder via either stalled ribosomes (ribosome display) or puromycin linkage (mRNA display) [159-162].

Directed evolution campaigns based on display technologies generally consist of four steps illustrated in **Figure 10**.

a. Diversification

The preparation of a combinatorial library is the first step. A combinatorial library is a collection of diverse DNA sequences. For the construction of a pool of binder variants, the number and position of randomized residues needs to be well defined in advance. The mutated residue positions can be contiguous or scattered along the coding sequence depending on their precise arrangement within the three-dimensional fold. The number of randomized residues and the set of allowed amino acid side chain exchanges determine the library diversity. The complexity of starting library, so-called naïve library, is an important factor that determines the success rate of the subsequently selection steps. The higher the complexity, the higher the number of possible variants. However, this does not mean that all targeted residues are actually mutated because of practical limits in molecular cloning techniques. The combinatorial library is usually

made by diversifying the parental DNA sequence of the scaffold at defined positions. These positions are randomized by introducing degenerate codons such as NNN (completely randomized including stop codon), NNS/NNK (partially randomized) where S = G or C, and K = G or T.

b. Translation

The library is then submitted to either in vitro transcription/translation (e.g. in ribosome display), or transformation into cells for in vivo transcription/translation (e.g. in yeast display) [158].

c. Selection

One of the critical factors of display techniques is library complexity. These are theoretical library sizes in different displays: phage (10^9), bacteria (10^{8-10}), yeast (10^7), ribosome (10^{13-14}) and mRNA (10^{13-14}). Since cell-free displays perform cell-free protein translation, there is no limit of cell transformation efficiency as in cell-based display techniques; hence, more diverse libraries can be generated in the former. Therefore, cell-free display may be suitable with naïve libraries for the first round of selection. Since library complexity is reduced after several rounds of selection, cell-based display technologies may be employed in the following steps [141, 163].

During the selection procedure, the immobilized target – typically protein ligands, are mixed with a pool of translated binders. Unbound scaffold mutants that do not bind strongly are washed away, and the bound binders are collected. The affinity between binders and the targets is enhanced by affinity maturation strategies: (1) Several rounds of selection, (2) Gradually increased stringency of selection such as tuning target concentration or association/dissociation time, as well as adding competitors with free ligands. (3) Randomization of coding sequence of selected binders by several options of mutagenesis such as error-prone PCR, site-directed mutagenesis, or gene shuffling [158].

d. Replication

The DNA of a pool of selected binders and/or random individual binders can be sequenced for alignment analysis, and the corresponding proteins are produced in larger scale for stability, affinity, and specificity measurements. The replicated genes serve as the initial source for subsequent rounds of library diversification [158].

3. Ribosome display

Ribosome display is an in vitro evolution technology suitable for the generation of binding proteins with high affinity for their targets. A prerequisite for the selection of binding proteins from combinatorial libraries is the coupling of genotype (mRNA) and phenotype (protein). This linkage is accomplished by an mRNA/ribosome/protein ternary complex that is formed during cell-free translation. To obtain such a complex, the 3' terminal DNA coding region of the library is genetically fused to a spacer lacking the stop codon. The stop codon in mRNA is responsible for recruitment of the release factors to stop translation. Therefore, when ribosome goes along

and decodes the mRNA to synthesize a polypeptide, without stop codon, the ribosomal complex does not receive the signal to be disassembled. Subsequently, the mRNA/ribosome/protein complex binds to the surface-immobilized target of interest via its nascent protein variant. Next, in the selection step, the unbound complexes are washed away, while the mRNA of bound complexes are collected and analyzed.

Cell-free translation systems can be obtained from bacteria extracts or purified translation components. Therefore, the naïve highly diverse library and the library after selection can be directly submitted for *in vitro* translation without an extra step of cell transformation. This minimizes the effect of reduced transformation efficiencies. The selection is performed by a surface-immobilized target coated on magnetic beads or microtiter plates for use on enzyme-linked immunosorbent assays (ELISA). After a binding event, the mRNA is eluted and reverse-transcribed to generate cDNA which continues to be used in the next round of selections or it is used directly to produce proteins for detailed biophysical characterization [164]. The ribosome display method is illustrated in **Figure 11**.

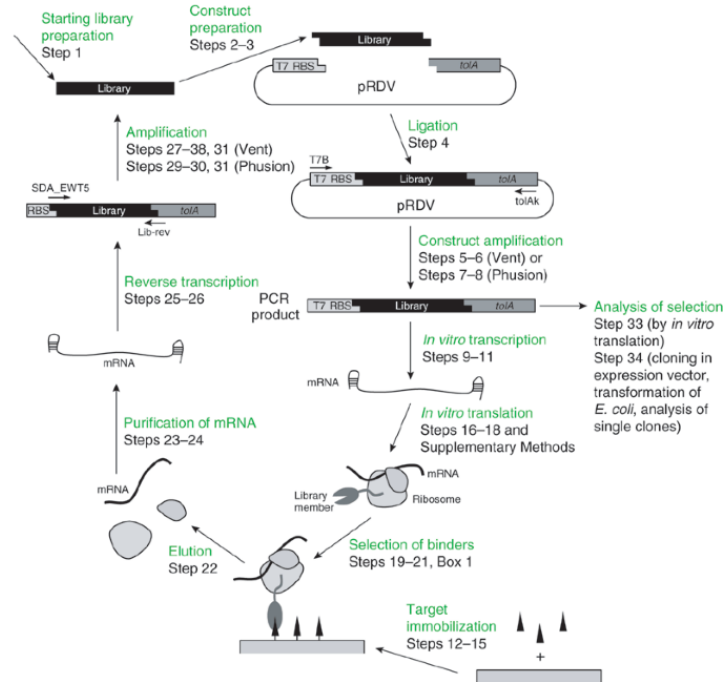


Figure 11: Illustration of ribosome display technology. Reprinted from [164] with permission, copyright © 2007, Springer Nature America, Inc.

E. Photoxenoprotein engineering

1. Introduction of photoxenoprotein engineering

Light is one of the most pervasive and persistent of environmental stimuli. Because of this, living matter possess a wide range of proteins, so-called photoreceptors or photosensory receptors, to sense and respond to electromagnetic radiation from the UV to the near infrared ranges [165]. These receptors convert light absorption into metabolic changes including alterations in gene expression, cell differentiation, molecular trafficking, etc. Overall, photosensory signaling underlies important biological processes like vision, circadian rhythms, and photosynthesis, to name but a few.

In order to study and manipulate biological processes with the same precision as nature, chemical biologists have developed a number of optical tools. The main advantages of using light to control protein activity are its non-invasiveness and spatiotemporal accuracy [166]. There are two major approaches to design photocontrol in proteins, both based on molecular engineering in response to light activation and deactivation. Fundamental to both approaches is using receiver molecules, chromophores, which absorb photons at a particular wavelength to subsequently change the structure and/or function of the protein in which they are embedded. Optogenetics (or hybrid protein optogenetics/optobiochemistry to be more precise [167]) involves fusing an intrinsically photosensitive protein to a target protein, while photoxenoprotein engineering involves the incorporation of a photosensitive group directly into the target protein [166, 168].

Protein functions such as ligand binding, catalysis, multimerization, conformational change, allosterism, and affinity to other proteins or nucleic acids often depend on a few critical residues. In photoxenoprotein engineering, these critical residues mediating protein activity are replaced with light-responsive non-canonical amino acids (ncAA), which render the protein light-responsive. As a result, proteins containing light-responsive ncAA, so-called photoxenoproteins, are generated. Upon exposure to light, the photoactive group of ncAA is transformed and thereby the protein activity is controlled [169].

2. Light-responsive non-canonical amino acids

Twenty genetically encoded canonical amino acids serve in nature as protein building blocks. Non-canonical amino acids are non-proteinogenic amino acids that are either found naturally in specific organisms or are synthetically made in a laboratory. There are three major types of light-responsive ncAAs including photocaged, photoswitchable, and photocrosslinkable ncAAs [169].

Photocaged ncAAs are powerful tools for transient blocking of binding interfaces. The photocaged ncAA can simply replace a residue that is essential for binding. Upon irradiation, the “cage” moiety of the ncAA is released (decaging or uncaging), through breaking of a covalent bond (photolysis), which sets free a canonical amino acid (or another ncAA) (**Figure 12A**). Thus,

after photo-induced cleavage of the light-sensitive protecting group, the biological activity of the protein is regained. The photodecaging reaction is irreversible. Most commonly used photocaged ncAAs are caged variants of tyrosine (containing nitrobenzyl, nitropiperonyl, or coumarin groups), or of lysine, cysteine, serine, histidine, glutamate, aspartate [166, 170]. One commonly used example is *ortho*-nitrobenzyl-tyrosine (NBY) [171] (**Figure 12A**).

Photoswitchable ncAAs can be used to achieve reversible manipulation of protein function. By inserting them close to the entrance of the binding site (or in an allosteric site), the ncAA can act as a lid blocking the entry of the ligand in one configuration. Illumination initiates bond isomerization (*trans/cis* or *E/Z*), and may be toggled in both directions by using suitable wavelengths, i.e., the photoisomerization reaction is reversible. Most available photoswitchable ncAAs are azobenzene derivatives of phenylalanine (e.g. AzoF) (**Figure 12B**) [172].

Photocrosslinkable ncAAs are inserted near a binding site so that protein-ligand interactions occur as in the wild-type system. Upon irradiation, the ncAA locks the ligand permanently onto the modified protein. To this end, protein–ligand or protein-protein interactions, especially those transient in nature, can be captured (“frozen”) and analyzed. The photocrosslinking reaction is also irreversible. Most photocrosslinkable ncAAs described so far are derivatives of phenylalanine and lysine containing benzophenones or diazirines, for example azibutyl-lysine (ABK) (**Figure 12C**) [173].

These photocontrollable ncAAs are utilized in a broad range of applications such as functional studies on proteins, tool development for in vivo fundamental studies, therapeutic approaches, production of challenging peptides and proteins and from biotechnology to industry [169]. The successful performance of photoxenoprotein engineering strongly relies on efficient incorporation of ncAAs to target proteins, often through genetic code expansion (GCE) technology [174, 175].

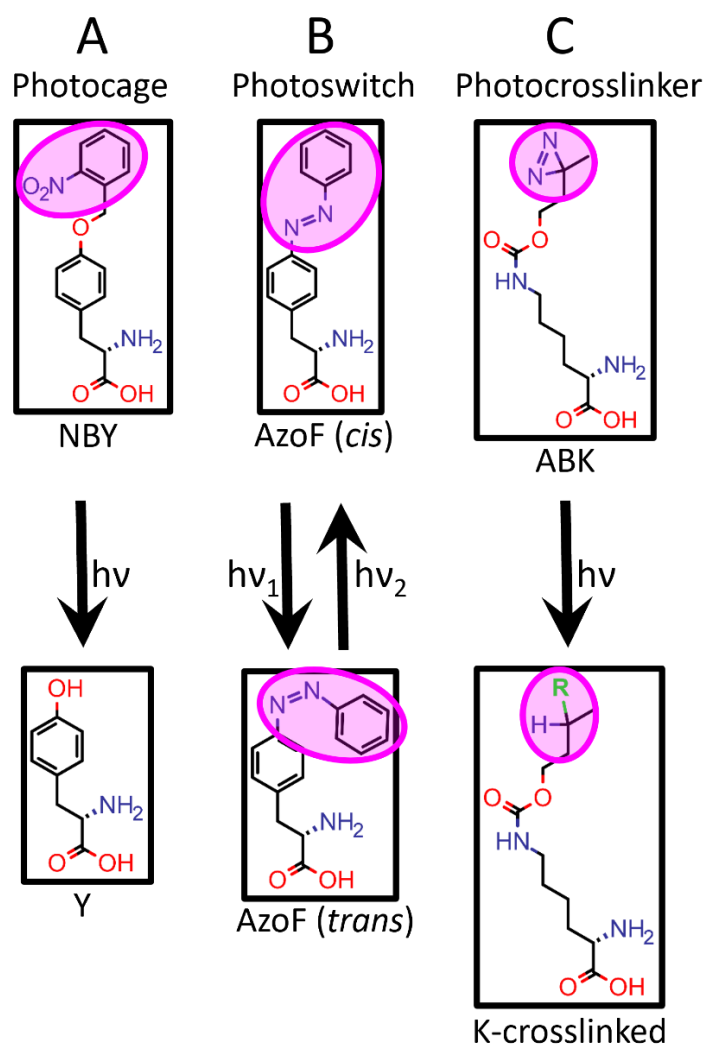


Figure 12: Examples of the three major classes of light-responsive ncAAs. (A) The photocaged ncAA NBY [171]. (B) The photoswitchable ncAA AzoF [172]. (C) The photocrosslinkable ncAA ABK [173]. NBY=(ortho-)nitrobenzyl-tyrosine, AzoF=azobenzene-phenylalanine, ABK=azibutyl-lysine, R=any atom. Figure created with Marvin Sketch.

3. Genetic code expansion technology

GCE refers to a set of tools for reprogramming the genetic code to enable the incorporation of more types of amino acid carrying desired properties, thus expanding the diversity and functionality of proteins [174-177]. Nowadays, more than 200 ncAAs can be incorporated into proteins by the GCE method for a wide range of applications. GCE includes several strategies to introduce ncAAs site-specifically into proteins in vivo e.g. stop codon suppression, sense codon reassignment, quadruplet decoding, nucleic acid alphabet expansion from four to six bases [178].

The standard or canonical genetic code consists of 64 codons, in which 61 are sense codons encoding 20 canonical amino acids and the remaining 3 are nonsense codons (UAA, UGA and

UAG) that do not encode any amino acid. Once the ribosome translation machinery encounters sense codons, they recruit the corresponding aminoacylated tRNAs, according to codon-anticodon interactions, to add an amino acid and prolong the polypeptide chain. The coupling of amino acid to tRNA is catalyzed by enzymes called aminoacyl-tRNA synthetases (aaRS), which hence constitute the “readers” of the genetic code. Once the ribosome encounters stop codons, they recruit release factors to terminate translation. Overall, four elements (ncAA, codon, tRNA and aminoacyl-tRNA synthetase) become targets for expanding the genetic code (**Figure 13**).

The stop-codon suppression strategy was developed during the eighties and is still nowadays the most utilized implementation of GCE [175, 179]. The stop codon suppressor tRNA_{CUA}, whose anticodon is modified to recognize the UAG (amber) stop codon, enables to insert its cognate non-canonical amino acid. In 1989, by using amber suppressor tRNA_{CUA}, which is chemically acylated by the specific ncAA, the ncAA was incorporated into protein *in vitro* [180]. Later in 2001, the orthogonal tyrosyl-tRNA-synthetase MjTyrRS/tRNA_{CUA} protein/RNA pair from *Methanocaldococcus jannaschii* (Mj) was added to the *E. coli* translation system enabling the host cells to incorporate non-standard amino acids into proteins. In principle, the native MjTyrRS charges tyrosine to its cognate tRNA_{CUA}, but not to any *E. coli* tRNAs. Evolution of this MjTyrRS/ tRNA_{CUA} pair by random mutagenesis allowed researchers to create a polyspecific synthetase that is able to incorporate other structurally similar ncAAs in response to the UAG codons in *E. coli* [175]. More and more orthogonal aaRS/tRNA_{CUA} pairs have been developed that expand the number of genetically encoded non-canonical amino acids in *E. coli* and other hosts, like mammalian cells. Among these ncAAs are several photocontrolled ncAA, including analogs of tyrosine [171, 181, 182], lysine [183, 184], cysteine [185, 186], serine [187], histidine [188], aspartate [189], glutamate [190, 191] and tryptophan [192].

Several difficulties complicating the use of the GCE method to introduce ncAAs are: (1) presence of truncated versions (along with the full-length protein) because of competition between suppressor tRNA and release factor for binding to the amber codon, (2) low expression yields (sometimes insufficient for subsequent studies) compared to wild-type [193].

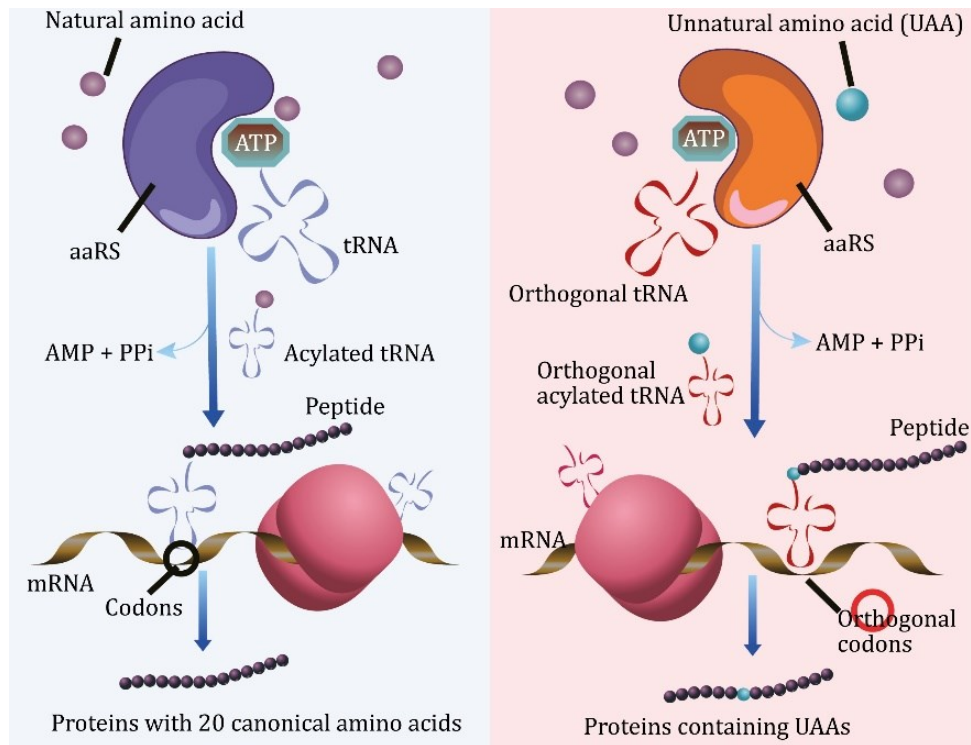


Figure 13: The genetic code expansion technology via stop codon suppression. Left panel, native translation system. Right panel, orthogonal translation system illustrating the four key elements: the aminoacyl-tRNA synthetase (aaRS), the tRNA, the codons, and the unnatural or non-canonical amino acids. Source: [194].

General steps for performing genetic code expansion via stop codon suppression:

- DNA construction: site-directed mutagenesis of the to-be-mutated residues on the target protein into the recoded codon (typically UAG), which must be different from the stop codon to end the translation (typically UAA or UGA).
- NcAA-protein expression: co-expression of plasmid encoding target protein with another plasmid containing the orthogonal translation system i.e. suppressor tRNA and aaRS specific for the ncAA to be incorporated. The NcAA itself must be usually supplemented to the cell culture medium.

II. GOALS

There are two general goals of this thesis: (I) studying the specificity of PPI and (II) modulation of the strength of PPI. As model proteins, I use human cytokines from the family of interleukin 10 and their receptors. The goals can be further divided as follows:

(I). Specificity of PPI

- A. To develop a general strategy to find new protein scaffolds and transform them into specific binders of human IL-10.
- B. To solve the structure of IFN- γ R2 by X-ray crystallography to help understand the molecular basis for IFN- γ specificity.

(II). Modulation of PPI

- C. To develop a photoxenoprotein engineering strategy based on genetically encoded photocaged ncAA and recombinantly expressed interleukin/receptor pairs to control the interaction between IL-24 and IL-20R2 by light.

III. RESULTS

A. Develop a general strategy to find new protein scaffolds and evolve a selected scaffold into an interleukin-10 binder.

Pham, P. N., M. Huličiak, L. Biedermannová, J. Černý, T. Charnavets, G. Fuertes, Š. Herynek, L. Kolářová, P. Kolenko, J. Pavlíček, J. Zahradník, P. Mikulecky and B. Schneider (2021). "[Protein Binder \(ProBi\) as a New Class of Structurally Robust Non-Antibody Protein Scaffold for Directed Evolution.](#)" *Viruses* 13(2).

My contribution: searching structures on PDB, sequencing analysis, cloning-expression-purification of wild type scaffold candidates and binder variants, ribosome display experiments, affinity measurement, protein crystallization.

Supplemented paper:

Kolenko, P., P. Mikulecky, **P. N. Pham**, M. Maly and B. Schneider (2023). "[Diffraction anisotropy and paired refinement: crystal structure of H33, a protein binder to interleukin 10.](#)" *J Appl Crystallogr* 56(Pt 4): 1261-1266.

My contribution: protein production and crystallization.

Summary

Antibodies have reached the drug market. However, nowadays is also well accepted that scaffold proteins with smaller size and simpler architecture provide structural plasticity to create high affinity binding interfaces against various medical ligands thereby complementing, and potentially surpassing, antibodies. Protein scaffolds have delivered promising results in different application fields. There are limitations to the use of the existing scaffolds such as patent monopoly, short half-life, immunogenicity, and low accessibility to the target area [141]. Therefore, we believe that mining novel scaffolds to enrich the current portfolio can provide more chances to yield safer and more potent drugs.

A compact structure with a rigid core and good tolerance to residue substitution without losing its structural integrity and stability is the most important attribute of a structurally robust protein scaffold. Other important properties are small size for good tissue penetration, thermostability, in vivo pharmacokinetics, and non-immunogenicity. All these features can be accounted for and thoroughly evaluated via a systematic stepwise selection procedure including search of suitable structures in the Protein Data Bank (PDB), a series of computational predictions of protein stability, and experimental screens. The most advanced scaffold would subsequently be mutagenized at defined residues to prepare a highly diversified DNA combinatorial library for in vitro directed evolution in combination with ribosome display to select the highest affinity variants against a medically important protein, IL-10.

The IL-10 is well known for its function as an immune repressor. Under physiological conditions, IL-10 is an important regulator of the immune system due to its immunosuppressive, anti-inflammatory, and stimulatory effects. Contrary to investigations suggesting its protective role in autoimmunity, allergy, and cancer, overexpression of IL-10 is associated with the pathogenesis of systemic lupus erythematosus, EBV-associated malignancies, and tuberculosis. Anti-IL-10 mAb were proved effective in clinical trials to treat patients with autoimmune disease SLE [63], suggesting that developing IL-10 antagonists is beneficial to control/treat cancers, EBV malignancies, and drug-resistant tuberculosis patients [63, 64, 67]. EBV not only uses its IL-10 homolog that binds IL-10 receptor to increase its proliferation and protect infected EBV-infected B-cells from immune recognition, but also enhances host IL-10 production to promote its survival and development. Therefore, neutralization/downregulation of IL-10 activity is a potential therapeutic approach.

Hereby, we present a step-by-step method starting from the selection of a novel protein scaffold, based on 4PSF, until the generation of variants with high affinity (~10-200 nM) against human IL-10. The workflow of these steps is shown in **Figure 14**.

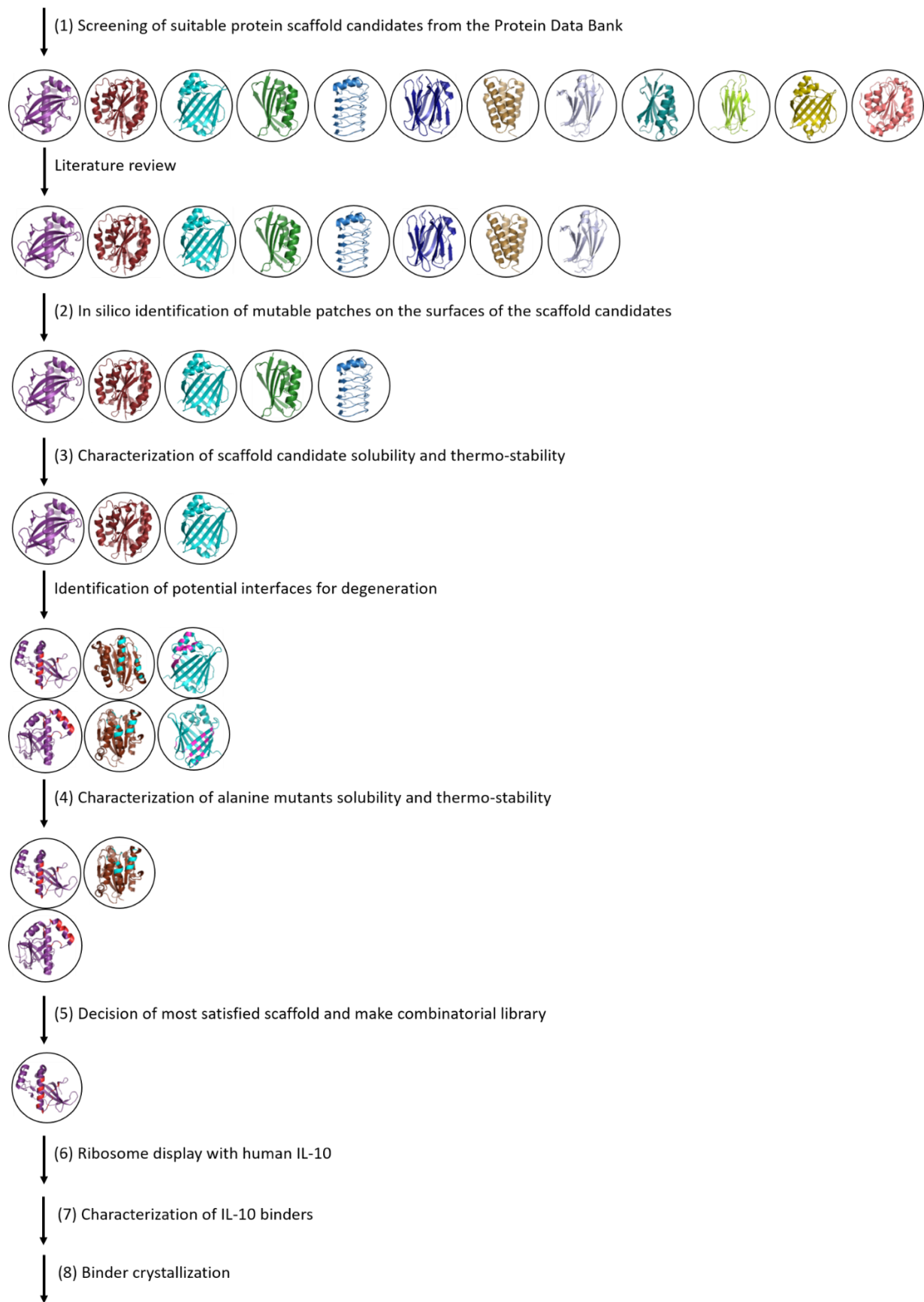
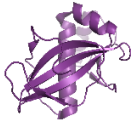

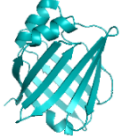
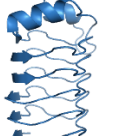





Figure 14: Workflow for the selection of novel scaffolds and their evolution.

1. Screening of suitable protein scaffold candidates from the PDB

The first step was to scan promising protein structures from the database of protein data bank (PDB). We set the selection criteria that protein candidates should possess high-resolution X-ray crystallography structure, molecular weight under 25 kDa, single domain, and expression in *E. coli*. The resulting hundreds of structures underwent further elimination based on literature review excluding structures with reported toxicity, ligand, cofactor addition for structure stabilization, or were previously unpublished as scaffold for directed evolution elsewhere. This process narrowed down the list to 12 final candidate structures listed in **table 3**.

PDB code	UniProt	Protein name	Structure	Resolution (Å)	Source organism	Size (kDa)	Expression host	Reference
4PSF	Q9NWS0	PIH1D1 N-terminal domain		1.58	<i>Homo sapiens</i>	15	<i>E. coli</i> BL21 (DE3)	[195]
1N3Y	P20702	Alpha-X beta2 integrin I domain		1.65	<i>Homo sapiens</i>	22	<i>E. coli</i> BL21 (DE3)	[196]
4I3B	P0DM59	Fluorescent protein UnaG wild type		1.20	<i>Anguilla japonica</i>	15	<i>E. coli</i> BL21 (DE3)	[197]
2F3L	B1WVN5	Luminal Rfr-domain protein		2.11	<i>Cyanothece sp. 51142</i>	20	<i>E. coli</i> BL21 (DE3)	[198]
1W2I	P84142	Acylphosphatase		1.50	<i>Pyrococcus horikoshii</i>	10	<i>E. coli</i> C41	[199]
4NBO	Q9HD15	Steroid receptor RNA activator protein carboxy-terminal domain		2.81	<i>Homo sapiens</i>	13	<i>E. coli</i> BL21 (DE3)	[200]
3APA	O60844	Human pancreatic secretory protein ZG16p		1.65	<i>Homo sapiens</i>	16	<i>E. coli</i> BL21 (DE3) codon +	[201]



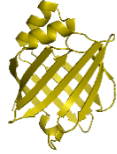


4IGI	E9PWQ3	Collagen VI alpha3 N5 domain		1.20	<i>Mus musculus</i>	22	<i>E. coli</i> BL21 Rosetta	[202]
2W4P	P07311	Human common-type acylphosphatase variant, A99G		1.70	<i>Homo sapiens</i>	11	<i>E. coli</i> C41	[203]
4LKT	Q01469	Human Epidermal Fatty Acid Binding Protein (FABP5)		2.57	<i>Homo sapiens</i>	15	<i>E. coli</i> BL21 (DE3)	[204]
4MJJ	Q14183	C2A domain of DOC2A		2	<i>Homo sapiens</i>	15	<i>E. coli</i>	To be published
4JOX	Q939T0	Cry34Ab1 protein		2.15	<i>Bacillus thuringiensis</i>	14	<i>E. coli</i>	To be published

Table 3: Structures and properties of twelve scaffolds. Protein structures created with ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197], 2F3L [198], 1W2I [199], 4NBO [200], 3APA [201], 4IGI [202], 2W4P [203], 4LKT [204], 4MJJ and 4JOX.

The 2W4P and 4LKT were excluded due to their structural similarity to candidates 1W2I and 4I3B. Moreover, 4MJJ and 4JOX were removed because of missing publications associated to PDB entries.

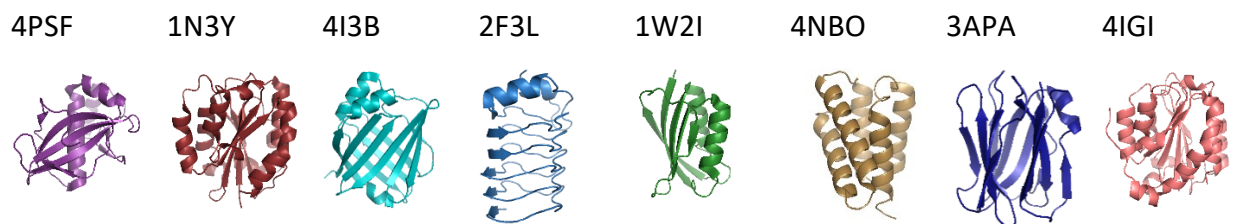


Figure 15: Eight selected scaffolds: 4PSF, 1N3Y, 4I3B, 2F3L, 1W2I, 4NBO, 3APA and 4IGI. Protein structures created with ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197], 2F3L [198], 1W2I [199], 4NBO [200], 3APA [201], 4IGI [202].

2. In silico identification of mutable patches on the surfaces of the scaffold candidates

The eight scaffolds (**Figure 15**) were analyzed to determine highly conserved residues, structurally essential residues, and poorly conserved residues. Both highly conserved residues and structurally essential residues, that are responsible for protein function and/or overall fold, must be preserved. Poorly conserved residues, which can in principle be substituted without overall structural change, would be targeted for constitution of the binding interface. The arrangement of conserved residues in space is expected to generate a rigid core structure that holds the poorly conserved region, which is more flexible and tolerant to residue substitutions. To select the residues to be mutated for the generation of a novel binding interface, we employed two strategies.

a. Multiple sequence alignment

Homologous sequences from different species were aligned to identify the percentage of conservation at each amino acid position. Highly conserved residues (>90%) were ruled out.

b. *In silico* saturation mutagenesis

Mutation scanning using the FoldX program was employed to evaluate the mutability of each residue. For each scaffold, each residue was substituted by the other 20 standard residues and calculated the corresponding free energy differences ($\Delta\Delta G$ in kcal/mol) between WT and mutants were calculated (kcal/mol). We subsequently calculated the mutability score of each residue as the percentage of mutations satisfying $\Delta\Delta G < 0$ (stability) or $0 < \Delta\Delta G < 0.5$ (slightly destabilizing).

Residues fulfilling conservation values lower than 90% and FoldX mutability scores larger than 50% were considered mutable residues. These selected residues were mapped on 3D structure to visualize whether they constitute mutable surface patches suitable for directed evolution. Since highly diverse library increases the probability of selecting desired binders against the target ligands and ribosome display can handle 10^{13} to 10^{14} variants, mutable surface patches containing 10 to 12 mutable residues were selected. Three scaffold candidates 4PSF, 1N3Y, 4I3B comprise of two mutable patches, while 2F3L, 1W2I consist of one mutable patch. Thus, both sets were selected. Other three scaffolds do not contain any surface patch were excluded (**Figure 16**).

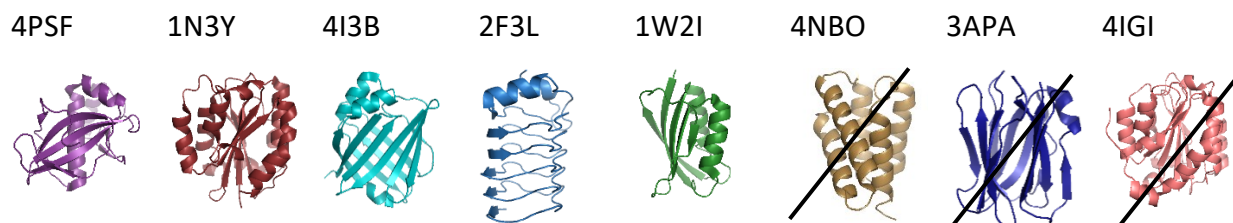


Figure 16: Five selected scaffolds (4PSF, 1N3Y, 4I3B, 2F3L and 1W2I) and three unselected scaffolds (4NBO, 3APA and 4IGI). Protein structures created with ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197], 2F3L [198], 1W2I [199], 4NBO [200], 3APA [201], 4IGI [202].

3. Characterization of the solubility and stability of scaffold candidates

We continued to check the solubility and purification simplicity of five scaffolds. Two constructs bearing an N-terminal His-tag and a C-terminal His-tag were made for each scaffold. Protein expression and solubility were done at four different temperatures (16 °C, 25 °C, 30 °C and 37 °C). Since 1W2I expression required a special bacteria strain, *E. coli* C43 (DE3), which were designed for toxic proteins, 1W2I was excluded. 4PSF, 1N3Y, 4I3B and 2F3L were highly soluble at 16 °C as N-terminal His-tag constructs, while C-terminal His-tag construct largely eliminated their expressibility. Moreover, while 4PSF, 1N3Y, 4I3B were highly pure after a two-step purification, 2F3L showed a contaminant on SDS-PAGE. Hence, we carried on with only 4PSF, 1N3Y, and 4I3B (**Figure 17**). Next, oligomerization of purified proteins was measured by dynamic light scattering (DLS), and the secondary structure and thermal stability were both determined by circular dichroism (CD) spectroscopy. All three scaffolds (4PSF, 1N3Y, 4I3B) were found mainly as monomers with the expected secondary structure content and high melting temperatures of 75 °C, 57 °C, and 47 °C, respectively.

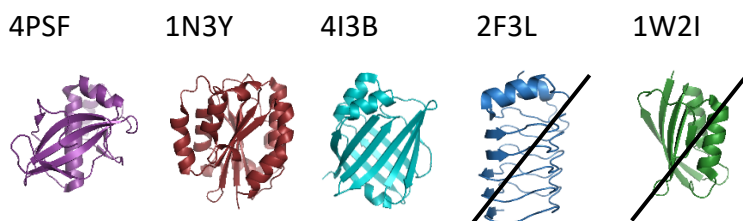


Figure 17: Three selected scaffolds (4PSF, 1N3Y and 4I3B) and two unselected scaffolds (2F3L and 1W2I). Protein structures created with ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197], 2F3L [198], 1W2I [199].

a. Potential interfaces for degeneration

The six mutable surface patches of the three scaffolds were highlighted (**Figure 18**). We named them based on the patch location near the N-terminus or C-terminus of the protein. 4PSF PatchN, 4PSF PatchC, 1N3Y PatchN, 1N3Y PatchC, 4I3B PatchN and 4I3B PatchC consisted of 10, 10, 10, 11, 12 and 11 mutable residues, respectively.

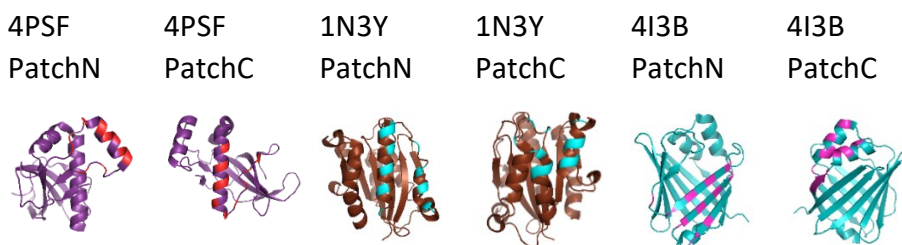


Figure 18: Three selected scaffolds (4PSF, 1N3Y and 4I3B) with two mutable surface patches (PatchN and PatchC) each (highlighted in red, cyan and pink respectively). Protein structures created by ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197].

4. Characterization of alanine mutants, solubility and thermostability

The tolerance to mutagenesis of the six scaffolds was validated by substituting all mutable residues on the surface patches to alanine. The expression, solubility, oligomerization and melting temperature of alanine mutants were measured in order to determine the “hit” mutants that behaved similarly as the wild type.

The Ala 4I3B PatchN and Ala 1N3Y PatchC aggregated into inclusion bodies. Low (10%) solubility compared to WT 4I3B was observed for Ala 4I3B PatchC. Three scaffolds (4PSF PatchN, 4PSF PatchC, 1N3Y PatchN) were selected (**Figure 19**) because they were as highly soluble as their WT, being also monodisperse and with acceptable high melting temperatures.

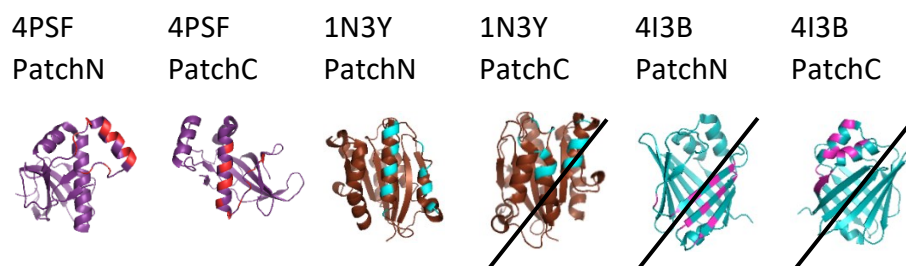


Figure 19: Three selected scaffolds (4PSF PatchN, 4PSF PatchC, 1N3Y PatchN) and three unselected scaffolds (1N3Y PatchC, 4I3B PatchN and 4I3B PatchC). Protein structures created with ChimeraX from PDB ID 4PSF [195], 1N3Y [196], 4I3B [197].

5. Decision of most satisfied scaffold and make combinatorial library

Sticking of protein on capillaries and sensor chips used in microscale thermophoresis (MST) and surface plasmon resonance (SPR) is problematic. Therefore, we check whether 4PSF, 1N3Y and their Ala mutants stick to these accessories. 1N3Y and 1N3Y PatchN were suitable for MST experiments but not for SPR. 4PSF, including its two mutants, were applicable for both MST and SPR. Considering these data together with the highest thermal stability (67 °C), we eventually decided 4PSF PatchC was our scaffold for directed evolution (**Figure 20**).

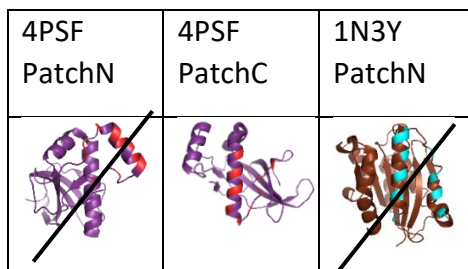


Figure 20: 4PSF PatchC was eventually chosen as the final scaffold for directed evolution. Protein structures created with ChimeraX from PDB ID 4PSF [195], and 1N3Y [196].

This 4PSF PatchC scaffold was called initially Probi. However, we renamed 57aBi due to trade mark of Probi. Ten mutable residues highlighted in red were randomized for DNA library construction. These residues were E110, T114, R117, E118, D122, N125, Q127, N129, E131, and R133 (numbering according to the 4PSF PDB entry).

QPGFCIKTNSSEGKVFINICHSPSIPPPADVTEELLQMLEEDQAGFRIPMSLGEPPHAELDAKGQGCTAYDVA
VNSDFYRRMQNSDFLR**ELVITIAREGLEDKYNLQLNPEWR**MMKNRPFMGSISQQNIR.

The DNA library was synthesized by GENEWIZ company by the degenerate NNK codons technology with estimation of 78 % library correctness, which was considered sufficient for further work.

6. Ribosome display with human IL-10

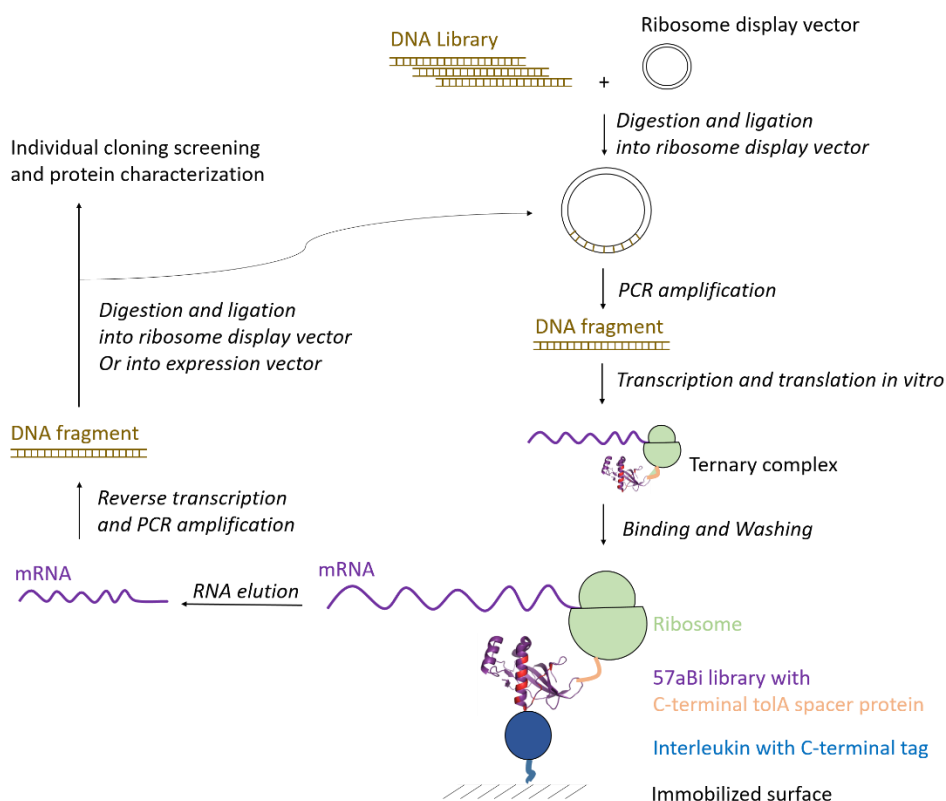


Figure 21: Scheme of ribosome display as set-up in our lab.

In ribosome display, the ternary complex (mRNA, ribosome, and binder), which is made by cell-free translation system, was incubated with surface-immobilized IL-10. After successful binding events and removal of non-specific binders, the mRNA is eluted, reversely transcribed to cDNA, and amplified via PCR technique. The DNA information of high affinity binders was then analyzed or another round of selection for affinity maturation to reach the desired affinity is added (**Figure 21**).

7. Characterization of IL-10 binders

After several rounds of ribosome display, the obtained binders, whose high affinity estimated by ELISA, were produced and their affinity were measured by MST. In initial preselection process, six variants bound unexpectedly background to BSA were acquired. Therefore, we optimized preselection process. As the result, two binders, named F5 and G3, interacted with IL-10 with high specificity and affinity, 6 nM and 208 nM, respectively (**Figure 22**). The melting temperature of these two binders were lower than WT protein, but still in the range of practical use. However, these binders were not able to inhibit signaling pathway of IL-10 and its receptors IL-10R2, checked by a competitive binding assay and phosphorylated STAT3 (pSTAT3) detection in WB.

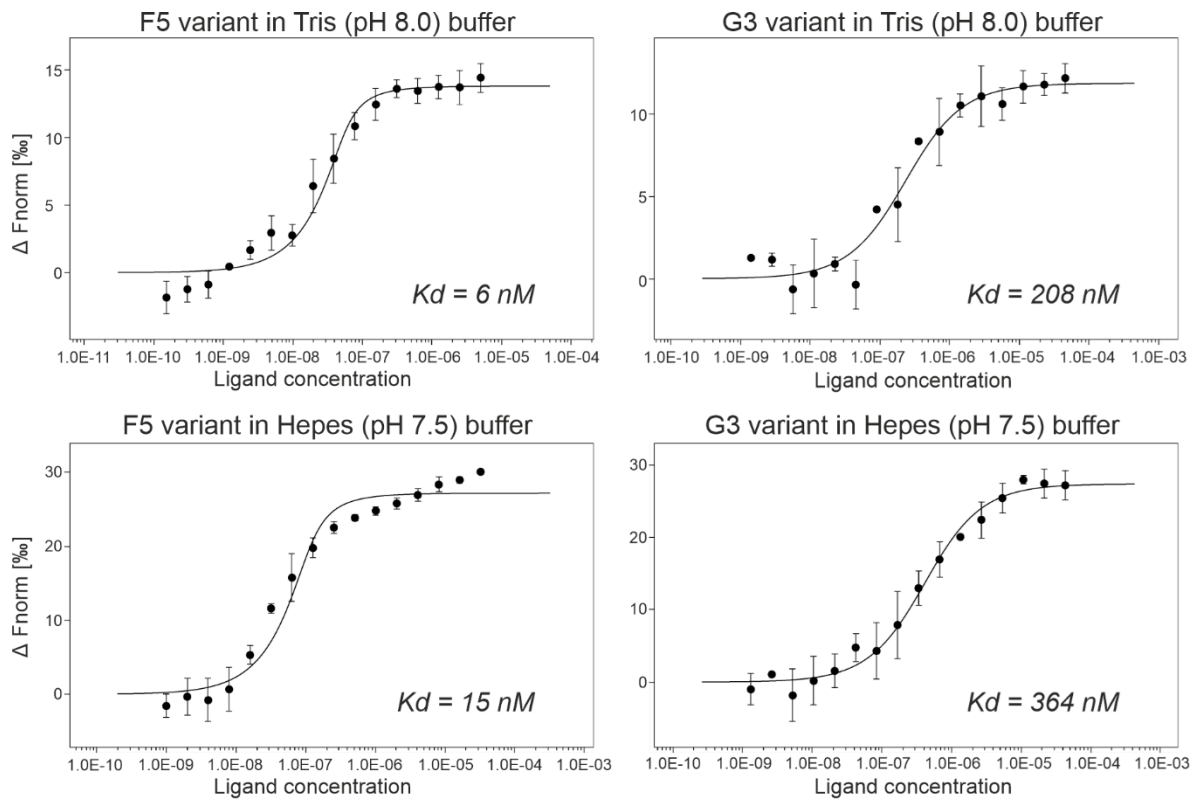
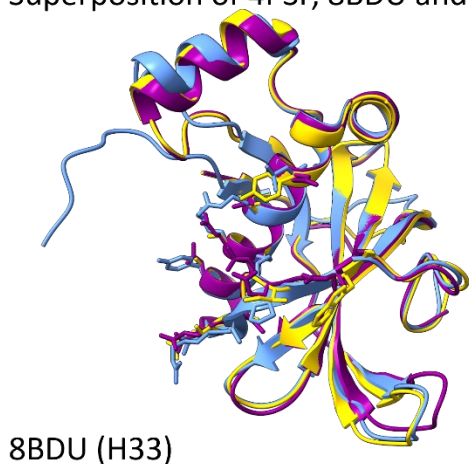


Figure 22: Determination of binding affinities (reported as dissociation constants, K_d) between evolved 57aBi binders and IL-10 by MSTs. Source: [7].

8. Binder crystallization

Six binder proteins arising from initial preselection process (in μM affinity range) were crystallized. Four crystals of six variants diffracted at a low resolution ranging from 6 to 8 Å. One binder formed crystal with high resolution 1.2 Å structure, deposited under the PDB entry ID 7AVC [7] and solved by the molecular replacement method using original 4PSF structure. The crystal structure of another binder at lower resolution (2.73 Å) was deposited under the PDB entry uID 8BDU [8] and solved by diffraction anisotropy and paired refinement method. In comparison, the overall structures of three proteins, one WT and two binders, are highly similar (**Figure 23**). This confirms that the fold of 5aBi scaffold is stable.

Superposition of 4PSF, 8BDU and 7AVC



4PSF (WT)

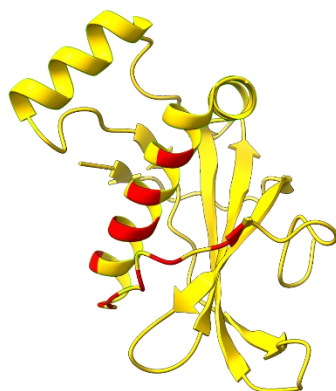


Figure 23: Comparison of protein structures between wild-type and two evolved binders. The superposition and individual 3D structures of three proteins demonstrates the stability of the fold: WT (violet), H33 (yellow) and J61 (blue) variant with PDB ID 4PSF [195], 8BDU [8] and 7AVC [7], respectively. Mutated residues are highlighted in red. Figure created with ChimeraX from the mentioned PDB IDs.

From the results collected through all steps of workflow, we confirmed that the randomization on 57aBi did not significantly affect the overall structure. Therefore, 57aBi is applicable for directed evolution by ribosome display.

B. Solving 3D protein structure of interferon gamma receptor 2

Mikulecky, P., J. Zahradnik, P. Kolenko, J. Cerny, T. Charnavets, L. Kolarova, I. Necasova, **P. N. Pham** and B. Schneider (2016). "[Crystal structure of human interferon-gamma receptor 2 reveals the structural basis for receptor specificity.](#)" Acta Crystallogr D Struct Biol 72(Pt 9): 1017-1025.

My contribution: assistance in protein crystallization and structure determination.

Summary

IFN- γ functions through binding to its cognate receptor, IFN- γ R2. Therefore, IFN- γ R2 plays important immune regulatory roles in innate and adaptive immunity against viral, bacterial and protozoal infections. The interaction between IFN- γ and IFN- γ R2 leads to activation of downstream signaling pathway but the structural basis of such interaction is elusive. A three-dimensional (3D) structure of IFN- γ R2, which has not been solved yet, would help to understand the signaling mechanism at the molecular level. There are three parts including:

- Biophysical characterization of the extracellular domain of human IFN- γ R2
- Structure determination of IFN- γ R2 by X-ray crystallography.
- Comparison of IFN- γ R2 to other cytokine receptors to find binding specificity.

1. Biophysical characterization of the extracellular domain of human interferon gamma receptor 2

N-linked glycosylation is critical for IFN- γ R2 production through stabilization of the protein fold and facilitation of the transport to and/or across the cellular membrane. Five of the six predicted glycosylation sites of IFN- γ R2 were confirmed by mass spectrometry and X-ray crystallography. There were no significant differences in melting temperature between the glycosylated and deglycosylated IFN- γ R2 proteins from CD spectroscopy and thermal shift assays. The secondary structure of the two proteins was highly similar. Moreover, removal of these glycosylation sites abolished protein expression.

Disulfide bonds are not critical for the stability of IFN- γ R2 fold. There are two S-S bonds identified by mass spectrometry. The melting temperature of glycosylated and deglycosylated forms of IFN- γ R2 treated with reducing agents is similar (~ 1 °C difference).

2. Structural features of interferon gamma receptor 2

The structure of IFN- γ R2 (UniProt entry P38484) was solved at 1.8 Å resolution (PDB ID 5EH1 [14]). It comprises an N-terminal D1 domain, composed of three β -strands stacked on a layer of four β -strands, and a C-terminal D2 domain, created by four β -strands arranged against four other β -strands, are connected by a short linker. The overall structure belongs to the immunoglobulin fold with fibronectin type III topology (**Figure 24**).

In further analysis, D1 domain contains a distinct structural motif of six stacked residues (Lys68, Trp74, Arg114, Trp126, Arg116 and His123) that contributes significantly to the overall stability of the whole domain (**Figure 24a**). The important role of tryptophan residues in stabilizing the fibronectin fold by stitching together two β -strands is supported by charged residues, especially arginine, in D1.

The structure reveals two sugar moieties attached to Asn110 and Asn137 (N-linked glycosylation), which sandwich the hydrophobic sidechain of Trp131 avoiding its exposure to the solvent (**Figure 24b**).

A unique structural feature of IFN- γ R2 is the specific positioning of an aromatic residue (Phe109) in the recognition epitope (**Figure 24c**).

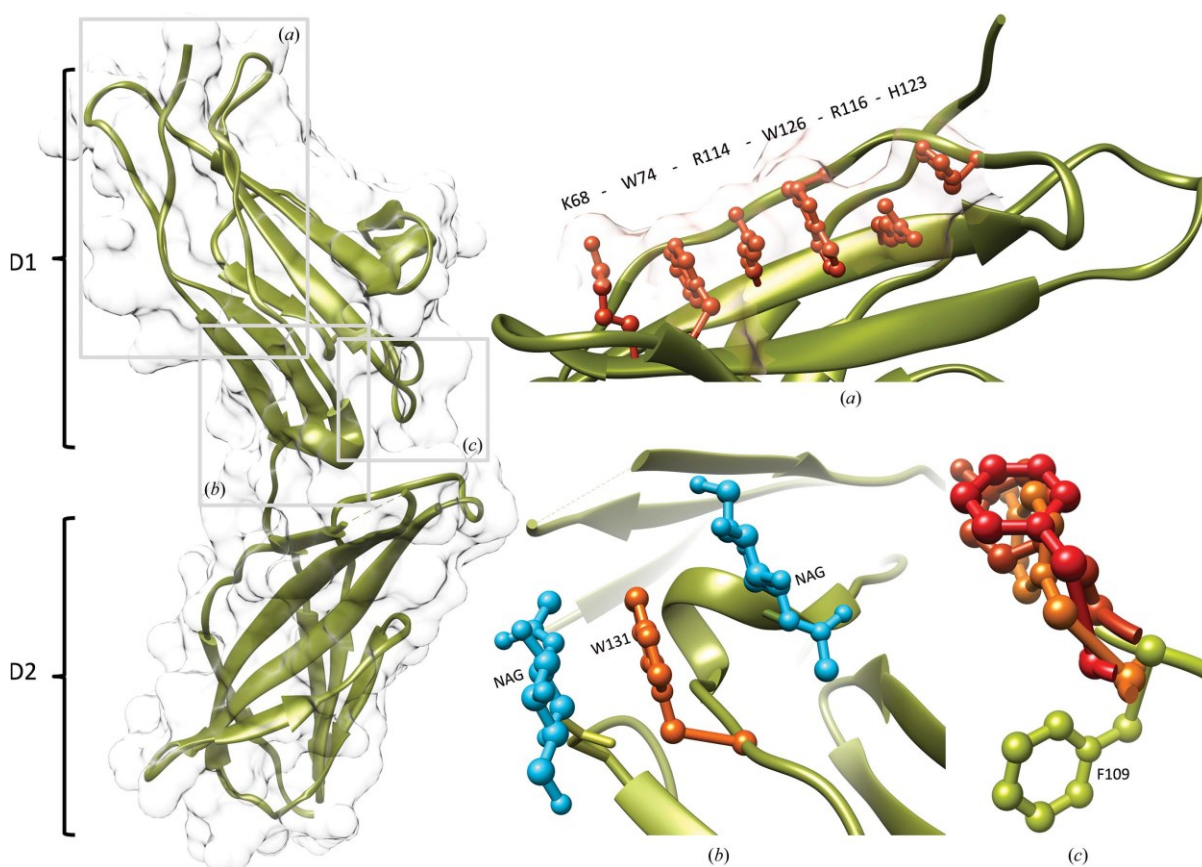


Figure 24: Structural characterization of IFN- γ R2 showing its two domains (D), D1 and D2 (PDB ID 5EH1 [14]) Source: [197].

(a) Residues K68, W74, R114, W126, R116 and H123 (in orange) of the D1 domain form a stacking motif on the IFN- γ R2 surface.

(b) Two N-Acetyl-d-glucosamine (NAG, in blue) sugars, bound to N110 and N137, sandwich W131 (in orange), thus decreasing its hydrophobicity.

(c) Superposition of the aromatic binding epitope between IFN- γ R2 (F109 in green) and other cytokine receptors (in orange, red) including IL-10R2 (Y82 PDB ID 3LQM [193]), gp130 (F169 PDB ID 1BQU [194]) and IL-2R2 (Y103 PDB ID 4GS7 [195]). There is no corresponding aromatic residue in IL-20R2 (PDB ID 4DOH [196]).

3. Structure and sequence alignments revealed important features of cytokine receptors.

Structure-based bioinformatics analysis in combination with multiple sequence alignment helped to determine the putative binding sites of IFN- γ R2 to IFN- γ and IFN- γ R1 (**Figure 25**).

PDB code	1fg9	5eh1	1j7v	3lqm	4doh	4doh	3dlq	3g9v	3og6	3se4	3se4	2puq
Chain ID	C	A	R	A	E	B	R	A	B	A	C	T
Receptor	IFN γ R1	IFN γ R2	IL10R1	IL10R2	IL20R1	IL20R2	IL22R	IL22BP	IL28R	IFN α R1	IFN α R2	TF
IFN γ R1	0.82	0.87	0.73	0.73	0.80	0.78	0.90	0.71	0.82	1.35	1.12	0.85
IFN γ R2	1.27	0.79	0.93	0.97	0.79	0.79	0.96	0.80	0.85	1.45	1.25	1.04
IL10R1	1.22	1.12	1.07	0.74	0.97	0.88	0.94	1.43	1.13	0.81	0.67	0.66
IL10R2	1.00	1.14	1.05	0.89	0.95	1.05	0.65	0.87	0.68	0.80	1.11	0.93
IL20R1	0.93	1.32	0.89	1.33	1.59	0.93	0.97	0.63	0.75	0.59	1.36	1.15
IL20R2	1.32	1.51	1.23	0.99	1.33	0.97	0.93	0.66	0.82	1.48	1.25	1.21
IL22R	1.20	1.49	1.09	0.86	1.04	1.16	4.19	1.20	1.19	0.87	0.98	1.03
IL22BP	1.32	1.48	0.94	1.20	1.18	1.30	1.17	0.93	1.02	1.16	1.11	0.96
IL28R	1.21	1.16	1.34	1.29	1.30	1.32	1.18	1.45	0.91	1.09	0.77	0.86
IFN α R1	1.16	0.96	1.31	1.33	1.33	0.98	1.75	1.26	1.59	1.77	0.94	1.20
IFN α R2	1.35	1.80	1.40	1.29	1.19	1.81	1.58	1.14	1.46	1.22	1.43	1.02
TF	1.38	1.16	1.23	1.26	1.44	1.04	1.16	1.40	1.43	1.16	1.81	0.86

Figure 25: Structural differences between D1 and D2 domains of 12 class 2 cytokine receptors quantified by their RMSD values. RMSD values smaller and larger than the average are colored in blue and red, respectively. Source: [14].

The residues of IFN- γ R2 highlighted in red had highly variable positions and were assumed to constitute the binding interface for IFN- γ and IFN- γ R1 (mapped on the structures in **Figure 26a**). Moreover, sequence alignment of IFN- γ R2 from distinct species revealed that the highest sequence variability occurs at positions coinciding with these structurally variable regions (**Figure 26b**).

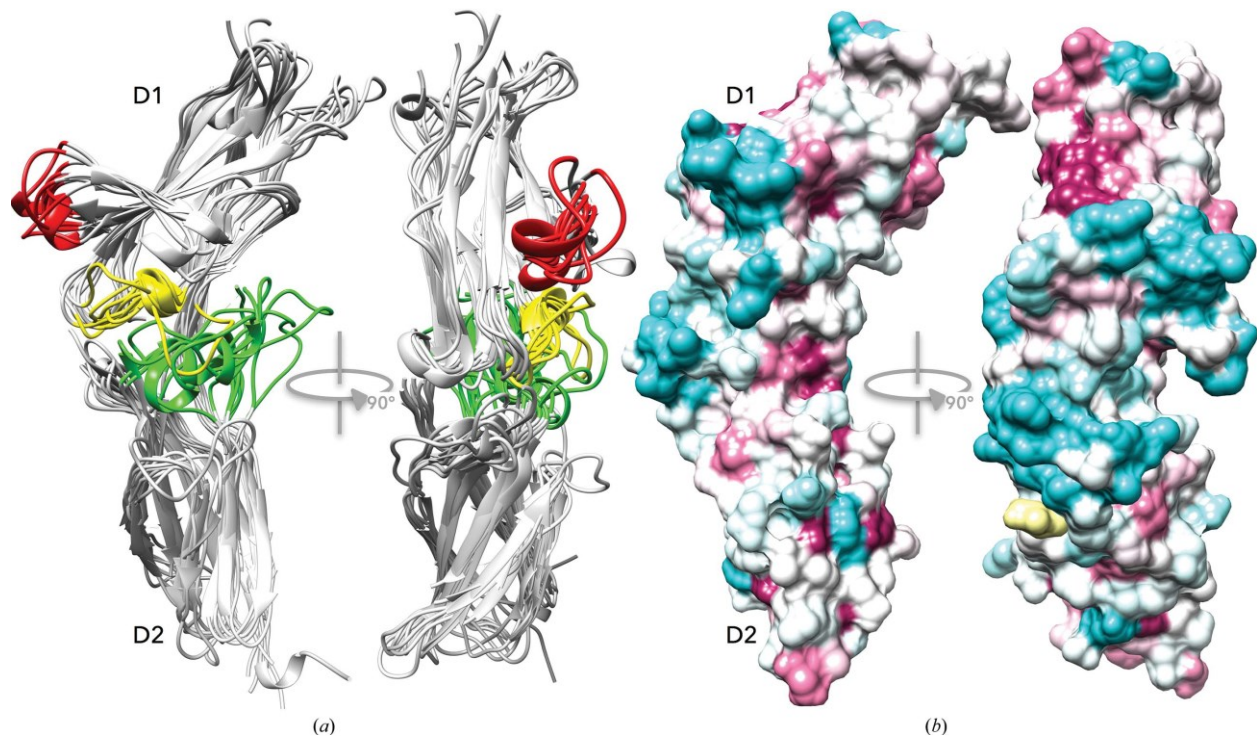


Figure 26: Structure and sequence alignments of IFN- γ R2 (PDB ID 5EH1). Source: [14].

(a) Structural alignment of D1-D2 domains from nine IFN- γ R2 variants. Two regions in D1 (red and yellow) and one region in D2 (green) are highly variable.

(b) Sequence alignment of IFN- γ R2 from 90 species with the sequentially most variable regions (cyan) and conserved regions (purple) illustrated on the surface of the IFN- γ R2 structure.

C. Modulation of the interaction between interleukin-24 and interleukin-20 receptor 2 by photoxenoprotein engineering

Pham, P. N., J. Zahradnik, L. Kolarova, B. Schneider and G. Fuertes (2023). "[Regulation of IL-24/IL-20R2 complex formation using photocaged tyrosines and UV light](#)." *Front Mol Biosci* 10: 1214235.

My contribution: photoxenoprotein design, cloning, expression and purification of parental IL, receptor and their ncAA variants, MST affinity measurement, cell signaling performance and protein detection by western blot.

Summary

Human IL-24 had been used as anti-tumor cytokine in Phase I/II clinical studies using virus-mediated gene delivery [99] and other delivery strategies [98, 103-105]. Tumor inhibition can be achieved either by overexpressing IL-24 inside the tumor or by overexpressing secreted IL-24 from distant cells based on the “bystander” activity. According to the latter effect, secreted IL-24 can “travel” and bind its receptors on the surface of tumor cells to trigger the Tyk2/Jak1 signaling pathway, leading to downstream activation of the transcription factor STAT3 via phosphorylation [86, 101]. Besides tumor treatment, IL-24 is also suggested as a therapeutic target for TB and IBD patients [109, 110].

Since many physiological and pathophysiological roles of IL-24 critically depend on its binding to cognate receptors, we believe that controlling such interaction between IL-24 and the shared receptor IL-20R2 could be beneficial for basic and applied research. Using light to switch ON/OFF cytokine-receptor interaction provides a high temporal and spatial resolution. Photoxenoprotein engineering, where “hot spot” residues on the target protein are replaced with light-responsive ncAAs such as NBY via GCE technology, is an approach commonly used for the temporal blockage of protein activity [205, 206]. NBY contains the “bulky” photocage nitrobenzyl, which is aimed at blocking the binding interface, and the tyrosine moiety. Upon illumination, the tyrosine-photocage bond is cleaved, leaving the canonical tyrosine on the protein and thus restoring binding. NBY has been applied successfully to block/unblock antigen-antibody interactions [207-212] but to the best of my knowledge has not been used to photocontrol non-antibody-based protein interactions yet.

We hereby demonstrated proof-of-concept for employing NBY and UV light to control the association between IL-24 and IL20-R2 (**Figure 27**). The presence of NBY on tyrosine 70 of IL-20R2, which is part of the binding interface with IL-24, successfully blocked the interaction. This was followed by its decaging, i.e. “cage” removal, upon illumination recovers the interaction. Existence of the ON/OFF binding was assessed by affinity measurements, from MST and yeast surface display, and cell signaling assays detecting phosphorylated STAT3.

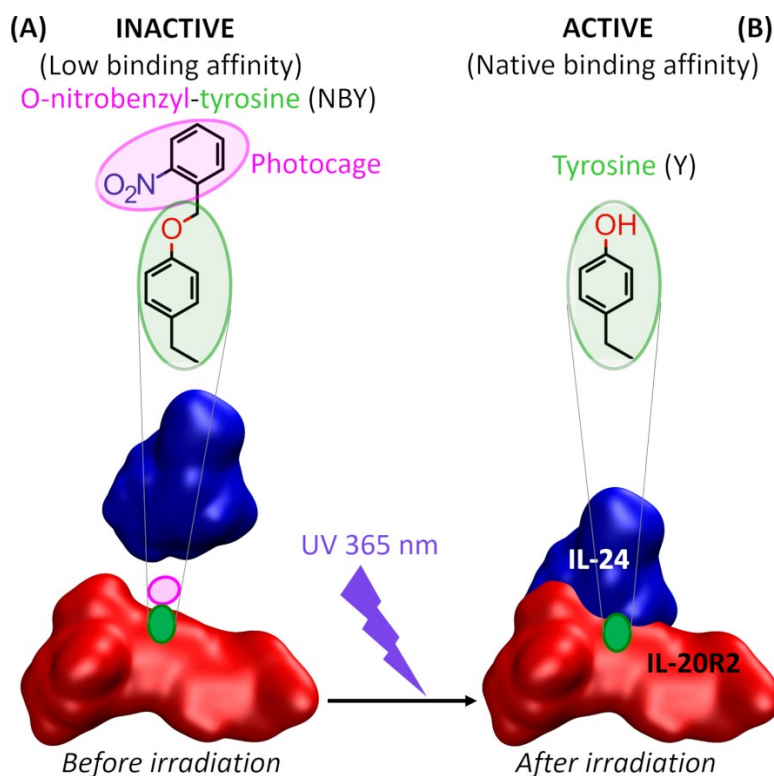


Figure 27: The combination of photocaged non-canonical amino acids, ortho-nitrobenzyl-tyrosine (NBY) residue, and UV light to control the interaction between IL-24 and IL-20R2. (A) In the absence of illumination, the presence of the cage at the interface blocks the association between the two partners. (B) After UV illumination, the cage is excised, the native tyrosine is generated, and the IL-24/IL-20R2 complex can be formed. Source [10].

1. Photoxenoprotein design

In order to achieve soluble expression in *E. coli*, we employed the PROSS algorithm to optimize the sequences of human IL-24 [9] and IL-20R2. The mutated IL-24 (renamed as IL-24B4) and IL-20R2 (renamed as IL-20R2D) sequences consisted of 29 and 23 mutations relative to their wild-type counterparts, respectively. According to the crystal structure of the ternary complex (IL-24/IL-20R2/IL-22R1) (PDB ID 6DF3) [79], these mutations were not located on the binding interface between IL-24 and IL-20R2.

Tyrosine is frequently found in the “hot spots” (regions of the amino acid residues on the PPIs interface that contribute to the binding-free energy) of binding interfaces [213]. We identified three tyrosine candidates at the interface of IL-24 and IL-20R2: two in IL-20R2 (Y70 and Y74) and one in IL-24 (Y204) (**Figure 28A**). The phenolic hydroxyl group of Y70 makes a hydrogen bond with the ϵ -amino group of lysine 135 in IL-24 (**Figure 28B**). The sidechain OH of Y74 is H-bonded to the backbone CO group of leucine 117 in IL-24 (**Figure 28C**). Y204 makes Van der Waals interactions with lysine 210 of IL-20R2 (**Figure 28D**). From these analyses, we prepared four constructs: three single mutants: IL-20R2D Y70NBY, IL-20R2D Y74NBY, and IL-24B4 Y204NBY; and one double mutant IL-20R2D Y70NBY/Y74NBY.

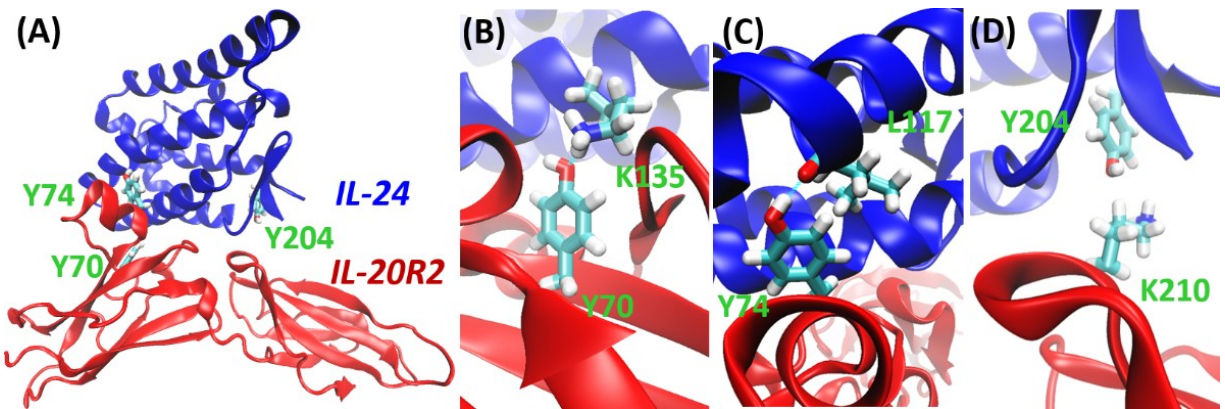


Figure 28. Selection of target residue positions for the installation of photocaged tyrosines. Source [204].

(A) Structure of the complex between IL-24 (blue) and IL-20R2 (red). The three target tyrosines (Y70 of IL-20R2, Y74 of IL-20R2 and Y204 of IL-24) participating in interleukin/receptor interactions are indicated in green.

(B) Interaction between Y70 (IL-20R2) and K135 (IL-24) via H-bonding.

(C) Interaction between Y74 (IL-20R2) and L117 (IL-24) via H-bonding.

(D) Interaction between Y204 (IL-24) and K210 (IL-20R2) via Van der Waals contacts.

Figure created with Visual Molecular Dynamics software from PDB ID 6DF3 [79].

2. Genetic code expansion for photoxenoprotein production

In order to obtain IL-24 and IL-20R2 photoxenoproteins, NBY photocaged ncAA and the evolved orthogonal tyrosyl-tRNA-synthetase (MjTyrRS*) and amber suppressor tRNA (tRNA_{CUA}) pair were employed in *E. coli* cells. Since the incorporation of NBY into cytokine and receptor is based on amber suppression, tyrosine codons at residue positions 70, 74 and 204 were replaced by TAG codons. The plasmids encoding for either IL-24B4 or IL-20R2D were co-transformed with plasmids encoding both aminoacyl-tRNA synthetase (aaRS) and tRNA_{CUA}. Bacterial media were supplemented with NBY.

3. Purification and characterization of photoxenoproteins

Because of the competition between tRNA_{CUA} and release factors, full-length proteins are mixed with truncated ones. Therefore, Strep tag was placed at the C-terminal end for protein purification to fish out the full-length proteins only. Size-exclusive chromatography was used to increase purity and to eliminate biotin from the affinity chromatography step. Six proteins including two parental IL-24B4, IL-20R2D and four mutants IL-20R2D Y70NBY, IL-20R2D Y74NBY, IL-24B4 Y204NBY, IL-20R2D Y70NBY/Y74NBY and IL-20R2D Y70NBY/Y74NBY were obtained in high purity as seen by well-defined single bands on SDS-PAGE. Mass spectrometry was used to confirm the presence of NBY in the mutants. The secondary structures of these mutants, characterized by circular dichroism (CD) spectroscopy, revealed that while IL-24B4 variants are mostly α -helical, IL-20R2D variants are richer in antiparallel β -sheets, in line with their crystal structures. Differential scanning fluorimetry indicated that IL-24 variants were more stable than IL-20R2D variants, with melting temperatures of $\sim 60^\circ\text{C}$ vs. $\sim 40^\circ\text{C}$, respectively.

4. Decaging efficiency upon UV irradiation

Decaging by UV irradiation is expected to release the nitrobenzyl moiety, i.e. the photocaged part, of NBY to both regenerate the canonical protein and restore binding. Six proteins in buffer solution, two parental and four mutants, were submitted to UV irradiation and their decaging efficiencies were assessed by mass spectrometry. The decaged mutant mass should be identical to the parental protein and lower than that of the non-decaged mutant mass by 135 Da (corresponding to the mass of the nitrobenzyl moiety) in the single-mutant case (270 Da in double-mutant case). MS data showed that all NBY-containing mutants were efficiently photolyzed by UV light, thereby regenerating the native tyrosine residue present at that position.

5. Monitoring photocontrol by affinity measurement in solution

a. Affinity measurement by MST

MST measurements, which are based on temperature-induced changes in fluorescence intensity, were done by using fluorescently labeled receptor IL-20R2 variants at a fixed concentration and mixing with varying concentrations of non-labeled parental IL-24B4. IL-24B4 Y204NBY was excluded because of insufficient protein yields. Two mutants, IL-20R2D Y74NBY and IL-20R2D Y70NBY/Y74NBY, bind to IL-24B4 with similar affinities (approximately 500 nM) in comparison to parental IL-20R2D (**Table 4**). Only the IL-20R2D Y70NBY variant showed low binding affinity (approximately 100 μ M) to its cytokine in the dark (**Figure 29A, Table 4**). From the MST result, IL-20R2D (-/+ UV) showed differential binding to affinities to IL-24B4 (**Figure 29A, Table 4**). This probably meant a side effect of UV on proteins. The binding affinities of IL-20R2D (+UV) and IL-20R2D Y70NBY (+UV) were similar (within 5%) suggesting that UV-light restores native-like binding affinities (**Figure 29A, Table 4**). We concluded that IL-20R2D Y70NBY was so far the most promising candidate for photocontrolling cytokine-receptor interactions in solution. We continued to confirm the behavior of IL-20R2D Y70NBY by another method, yeast surface display.

Partner 1	Partner 2	MST	Yeast display	Cell signaling (pSTAT3)
IL-24B4	IL-20R2D	K_d (nM)	K_d (nM)	EC_{50} (nM)
IL-24B4	IL-20R2D	630±1400.3	5.0±0.6±0.3	0.28±0.07
IL-24B4 Y204NBY	IL-20R2D	N.P.	N.D.	0.9±0.5
IL-24B4 Y204NBY + UV	IL-20R2D	N.P.	N.D.	0.76±0.24
IL-24B4	IL-20R2D + UV	3700±400	9.6±1.6	N.P.
IL-24B4	IL-20R2D Y70NBY	>94000	>3400.3	N.P.
IL-24B4	IL-20R2D Y70NBY + UV	3300±600	7.6±0.4	N.P.
IL-24B4	IL-20R2D Y74NBY	430±0.30	N.D.±0.3	N.P.
IL-24B4	IL-20R2D Y70NBY/Y74NBY	570±60	N.D	N.P.

Table 4: Dissociation constant (K_d) values (measured from MST and yeast display assays with purified proteins in solution) and half-maximum effective concentrations (EC_{50}) values (measured from pSTAT3 production in cell signaling assays) of IL-24B4 and IL-20R2D variants. Source [10].

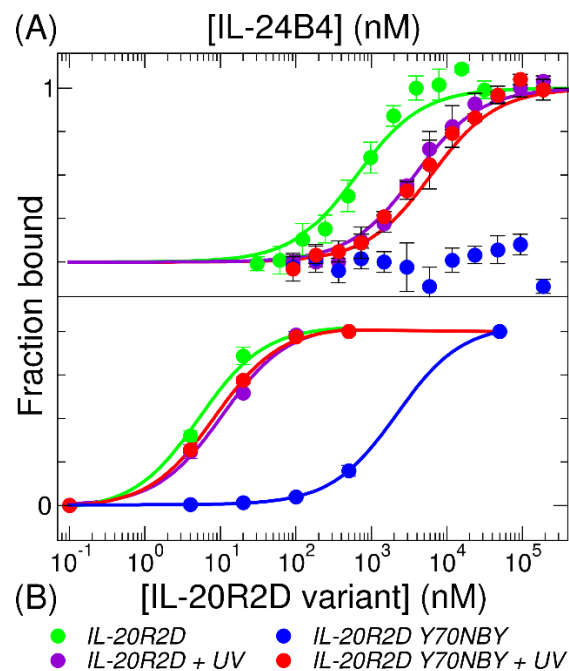


Figure 29. Quantification of binding affinities by two different methods: MST and yeast surface display. Solid lines in panels (A) and (B) are fits to retrieve the dissociation constants (K_d) values shown in **Table 4**. Source: [10].

(A) MST experiments.

(B) Yeast surface display experiments.

b. Affinity measurement by yeast surface display

In yeast surface display, the set-up is opposite to MST. The yeast, which expressed IL-24B4 (fixed concentration) fused to a fluorescent protein on its surface, was incubated with different concentrations of fluorescently labeled receptors IL-20R2D (-/+UV) and IL-20R2D Y70NBY (-

/+UV). Binding events were detected by fluorescence from each binding partner and then analyzed by flow cytometry.

In agreement with MST results, yeast display showed low levels of complex formation between IL-20R2D Y70NBY and IL-24B4 in the absence or UV illumination. High binding affinities, similar to the parental IL-20R2D, were observed upon UV irradiation and subsequent decaging (**Figure 29B**). We noticed that the binding affinities estimated by MST in diluted protein solution were lower (approximately 100-fold) than those determined by yeast display in crowded membrane environment (**Figure 29, Table 4**), most likely due to avidity (multiple interaction events) effects. In summary, from both methods, we concluded that IL-20R2D Y70NBY is a potential photoxenoprotein to control the interaction with IL-24 by UV light in solution.

6. Monitoring photocontrol by detection of activated signaling pathway

The effect of the presence/absence of interaction between IL-24 and IL-20R2 can be proved via the presence/absence of phosphorylated STAT3 (pSTAT3), respectively (**Figure 30**), which is a well-known marker for intracellular signaling dependent on IL-24 and other cytokines. The parental IL-24B4 with/without UV decaging (-/+ UV) and the IL-24B4 Y204NBY (-/+ UV) were added to the culture medium of HeLa cells for 30 minutes to trigger signaling, then cells were harvested and proteins were extracted. Next, pSTAT3 was detected by anti-pSTAT3 mAbs via SDS-PAGE and Western Blot (WB). Since the limited expression levels of endogenous receptors IL-20R2 and IL-22R1 in HeLa cells leads to a weak signal of pSTAT3, transient transfection with plasmids encoding the receptors is necessary to achieve amounts of pSTAT3 that can be quantified accurately.

In negative control experiments with transfected cells alone, little pSTAT3 was present in cells (background). Two different concentrations of IL-24 variants were added to saturate the pSTAT3 amount and detection. pSTAT3 was detected as a thick band in response to the addition of parental IL-24B4 (-/+ UV), showing that this engineered IL-24 version was able to activate the JAK/STAT signaling pathway and that UV illumination did not significantly affect IL-24 function. pSTAT3 levels in response to IL-24B4 Y204NBY were somewhat higher in the presence than in the absence of UV illumination. It meant this nCAA mutant blocks to some extent the binding to its receptor, and photocage removal partially restored the interaction. The differential response to UV excitation was small, but it was initial evidence to show that IL-24B4 Y204NBY could be a promising candidate to photocontrol protein interactions in a cellular environment.

We have not tested IL-20R2D Y70NBY in cell signaling assays due to the background levels of pSTAT3 production upon IL-20R2D addition. However, these experiments might be optimized in the future.

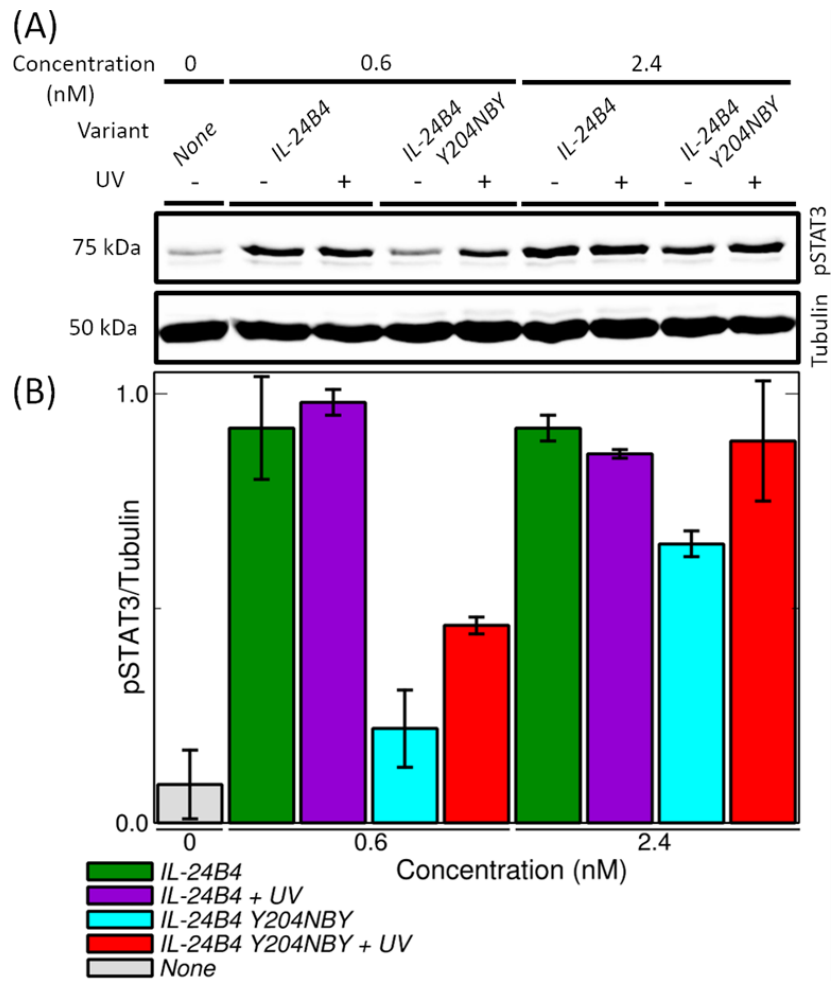


Figure 30: Signaling assays in HeLa cells expressing IL-20R2 and IL-22R1. Source [204].

(A) Production of phosphorylated STAT3, detected by antibodies anti-pSTAT3, upon addition of two interleukin-24 variants (before/after UV decaging at $\lambda = 365$ nm, for 5 min, at 100 mW).

(B) Quantification of pSTAT3 production by Western blot (normalized by tubulin) as a function of the added interleukin-24 variants (before/after UV decaging) and concentrations. The bars indicate the mean and standard deviation of two independent experiments.

IV. DISCUSSION

A. Development of 57aBi scaffold and evolution of 57aBi-based binders.

An intriguing and so far unexploited feature of our main scaffold, so-called 57aBi, is that it contains two mutable interfaces, N- and C-terminal, that may evolved to interact simultaneously with two different partners. These double binders could mimic the interaction mechanism of natural cytokines, which is based on the formation of a ternary complex with two receptors [11]. Furthermore, this approach would lead to the creation of bispecific binders, which has been achieved to some extent with antibodies [214] but are typically challenging to design [215]. Moreover, 57aBi-derived bispecific binders could work as “synthekines” that bind to non-natural receptor hetero-dimers to activate new unexplored cellular responses [216].

Our second developed scaffold, the 1N3Y PatchN, designed from protein alpha-X beta2 integrin I domain [196], could also be employed for binder development in the near future [16].

Protein structures that failed our stringent selection process, for instance 4I3B and 1N3Y PatchC lacked expression of Ala variants, could still be reconsidered as scaffolds. We noticed they contain 11 and 12 mutable residues, respectively, compared to 10 residues in the case of 4PSF and 1N3Y PatchN. By reducing the number of mutable residues, these proteins might pass the selection criteria and become novel scaffolds.

Ribosome display technology was prioritized due to the high complexity of scaffold library. However, other display technologies could also be included in our pipeline. Specifically, the yeast display technology, which provides certain advantages such as high-throughput detection by fluorescence-activated cell sorting (FACS), may be utilized for binder selection, affinity maturation and improvement of specific features [217].

Even though we do not expect 57aBi-derived binders to be immunogenic, since the scaffold originates from a human protein [195], the immunogenicity of binders has to be checked on a case-by-case basis for use in medical purposes. If binders would turn immunogenic and hence unsuitable for therapeutical applications, they could still be used for molecular imaging.

Binders G3 and F5 did not showed inhibitory ability by sequestering IL-10 and preventing activation of the signaling pathway. In order to identify truly inhibitory binders, there are several approaches that could be employed such as testing other binders, using different cell types (THP-1 or U937), or trying different signaling assays besides pSTAT3 detection.

The cross-reactivity of IL-10 binders to other cytokines and other unrelated proteins will be checked.

B. The structure of interferon gamma receptor 2.

Our structure of isolated IFN- γ R2 promotes the exploration of structure of the full ternary complex, thus rendering a better understanding of the functioning of this immunity cascade. Indeed, the crystal structure of the ternary signaling complex (PDB ID 6E3K [128]) was eventually solved by molecular replacement using our deposited model (PDB ID 5EH1) [128]. The structural superposition of this IFN- γ R2 in the ternary complex and our IFN- γ R2 shows high similarity (**Figure 31**).

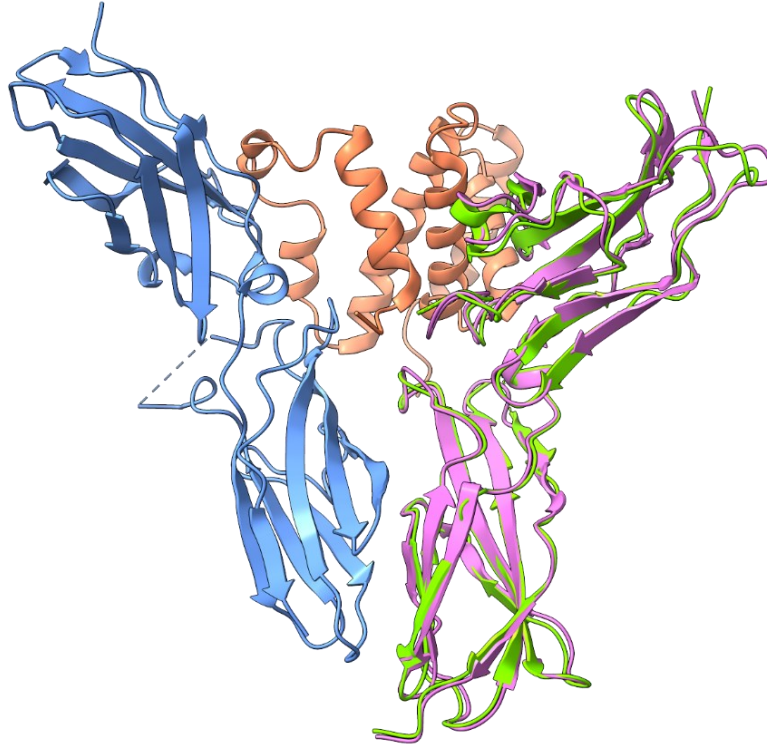


Figure 31: Overlaid cartoons of IFN- γ ternary complex (IFN- γ in orange, IFN- γ R1 in blue and IFN- γ R2 in pink, coordinates taken from [128]) and our IFN- γ R2 in green (coordinates taken from [14]). The structure reveals that IFN- γ R2 makes extensive contacts with both the IFN- γ dimer and with IFN- γ R1. Figure created with ChimeraX from PDB ID 6E3K (ternary complex) and 5EH1 (free IFN- γ R2).

Interestingly, a novel splice site mutation in patients with primary immunodeficiency has been found recently [218]. The mutation caused deletion of three amino acids in one of the variable loops of IFN- γ R2 that we predicted to control the binding specificity for IFN- γ .

C. Modulation of the interaction between interleukin-24 and interleukin-20 receptor 2 by light.

NcAA incorporation by GCE may cause a decrease of protein yield [219]. Therefore, it is important that the parental proteins can be produced in decent amounts. The bioinformatics-designed IL-24B4 and IL-20R2D, containing a large number of mutations compared to the WT, are suitable as parental proteins for this ncAA photoxenoprotein engineering study because they can be produced in *E. coli* in sufficient quantity and quality.

The IL-20R2D Y70NBY, one of all four produced variants, was a “hit” in the sense that it was able to modulate cytokine/receptor interactions as examined by MST and yeast surface display. IL-20R2D Y70NBY showed low affinity for IL-24B4 in the dark, and regained native-like (nM) affinity after UV light activation. This result agreed with several antibody studies using antibody fragments, scFv, nanobodies and intrabodies, which showed that a single NBY residue at a critical position is sufficient to diminish antigen-antibody interactions, while binding is restored upon light exposure [207-210].

The parental IL-24B4 and caged/decaged IL-24B4 Y204 variants showed minimal albeit reproducible differences in cell signaling assays. The apparent higher efficiency of Y70 in blocking heterocomplex assembly compared to Y204 might be due to its location at the core of the interface [79].

While the IL-20R2D Y70NBY clearly reduced the binding affinity for IL-24B4, the IL-20R2D Y74NBY and the double mutant IL-20R2D Y70NBY/Y74NBY did not significantly alter PPI. This puzzling result is in agreement with the non-additive effect of mutations on protein activity, known as epistasis [220, 221].

Instead of using a soluble version of IL-20R2D, one could directly express the full-length receptor containing NBY on the surface of HeLa cells. This can be done by employing a NBY-specific orthogonal translation system derived from *Methanosarcina barkeri* in future studies [181]. Alternatively, a competition cell-based assay could be performed in which the soluble photocaged receptor would interfere the binding of the cytokine to the cell-surface receptor. However, in the negative control having soluble IL-20R2D without cytokine, cellular pSTAT3 was unexpectedly detected. This problem can be circumvented by using universal reporter cell lines and/or reporter assays based on gene expression [222, 223] instead of pSTAT3 detection by WB.

Moreover, photocaged ncAA permit a one-time activation of PPI, which may not always be desirable. For instance, sustained cytokine-dependent JAK/STAT may drive excessive transcription of genes involved in inflammation, angiogenesis and survival [224]. Alternatively, photoswitchable ncAA [172, 225, 226] may enable bidirectional photocontrol of cytokine-receptor interactions for more demanding applications where reversibility is essential. Photoswitchable cytokines could be activated with light of a given wavelength (triggering complex formation with the paired receptor and subsequent phosphorylation of STAT3), and deactivated with light of another color that would switch off the JAK/STAT cascade.

V. CONCLUSIONS

The conclusions of the present Ph.D. thesis can be summarized as follows:

(I). Specificity of PPI

A. Selection and evolution of novel scaffolds specific for IL-10.

A systematic procedure based on three pillars (bioinformatics analysis of protein structures deposited on PDB, theoretical predictions and experimental examinations) was developed to eventually select two novel protein scaffolds: the human PIH1 domain-containing protein 1 (PDB ID 4PSF [195]), and the alpha-X beta2 integrin I domain (PDB ID 1N3Y [196]).

One of the two mutable interfaces of 4PSF, designated as 57aBi, was trained through ribosome display to bind specifically to the medically important human IL-10 with nanomolar affinity.

B. The crystal structure of the human IFN- γ R2 reveals the structural basis for receptor specificity.

According to structure and sequence alignments of related receptors, three structurally highly variable regions most likely bring about binding specificity for IFN- γ .

(II). Modulation of PPI

C. Modulation of the interaction between the IL-24 and IL-20R2 by photoxenoprotein engineering.

A photocaged tyrosine residue (NBY) introduced by GCE at a critical residue position in IL-20R2, Y70, diminished the formation of IL-24/IL-20R2 complex in solution. Upon UV illumination, the hetero-complex was assembled with nanomolar affinity.

A photocaged tyrosine residue (NBY) introduced by GCE at a critical residue position in IL-24, Y204, disrupted to a certain extent the JAK/STAT signaling pathway in cells that could be partially reactivated upon UV illumination.

VI. REFERENCES

1. Lu, H., et al., *Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials*. Signal Transduct Target Ther, 2020. **5**(1): p. 213.
2. Xie, X., et al., *Recent advances in targeting the "undruggable" proteins: from drug discovery to clinical trials*. Signal Transduct Target Ther, 2023. **8**(1): p. 335.
3. Fuertes, G., L. Nevola, and S. Esteban-Martín, *Perspectives on drug discovery strategies based on IDPs*, in *Intrinsically Disordered Proteins*. 2019. p. 275-327.
4. Commins, S.P., L. Borish, and J.W. Steinke, *Immunologic messenger molecules: cytokines, interferons, and chemokines*. J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S53-72.
5. Mollazadeh, H., et al., *Immune modulation by curcumin: The role of interleukin-10*. Crit Rev Food Sci Nutr, 2019. **59**(1): p. 89-101.
6. Akdis, M., et al., *Interleukins, from 1 to 37, and interferon- γ : Receptors, functions, and roles in diseases*. Journal of Allergy and Clinical Immunology, 2011. **127**(3): p. 701-721.e70.
7. Pham, P.N., et al., *Protein Binder (ProBi) as a New Class of Structurally Robust Non-Antibody Protein Scaffold for Directed Evolution*. Viruses, 2021. **13**(2).
8. Kolenko, P., et al., *Diffraction anisotropy and paired refinement: crystal structure of H33, a protein binder to interleukin 10*. J Appl Crystallogr, 2023. **56**(Pt 4): p. 1261-1266.
9. Zahradník, J., et al., *Flexible regions govern promiscuous binding of IL-24 to receptors IL-20R1 and IL-22R1*. The FEBS Journal, 2019. **286**(19): p. 3858-3873.
10. Pham, P.N., et al., *Regulation of IL-24/IL-20R2 complex formation using photocaged tyrosines and UV light*. Front Mol Biosci, 2023. **10**: p. 1214235.
11. Kolarova, L., et al., *De novo developed protein binders mimicking Interferon lambda signaling*. FEBS J, 2022. **289**(9): p. 2672-2684.
12. Mikulecky, P., et al., *Increasing affinity of interferon-gamma receptor 1 to interferon-gamma by computer-aided design*. Biomed Res Int, 2013. **2013**: p. 752514.
13. Cerny, J., et al., *Redesigning protein cavities as a strategy for increasing affinity in protein-protein interaction: interferon- gamma receptor 1 as a model*. Biomed Res Int, 2015. **2015**: p. 716945.
14. Mikulecky, P., et al., *Crystal structure of human interferon-gamma receptor 2 reveals the structural basis for receptor specificity*. Acta Crystallogr D Struct Biol, 2016. **72**(Pt 9): p. 1017-25.
15. Zahradnik, J., et al., *Interferons type II and their receptors R1 and R2 in fish species: Evolution, structure, and function*. Fish Shellfish Immunol, 2018. **79**: p. 140-152.
16. Huliciak, M., et al., *Combined in vitro and cell-based selection display method producing specific binders against IL-9 receptor in high yields*. FEBS J, 2023. **290**(11): p. 2993-3005.
17. Wei, H., et al., *Interleukin-10 Family Cytokines Immunobiology and Structure*. Adv Exp Med Biol, 2019. **1172**: p. 79-96.
18. Ouyang, W. and A. O'Garra, *IL-10 Family Cytokines IL-10 and IL-22: from Basic Science to Clinical Translation*. Immunity, 2019. **50**(4): p. 871-891.
19. Zdanov, A., *Structural analysis of cytokines comprising the IL-10 family*. Cytokine Growth Factor Rev, 2010. **21**(5): p. 325-30.
20. Chen, J., R.R. Caspi, and W.P. Chong, *IL-20 receptor cytokines in autoimmune diseases*. J Leukoc Biol, 2018. **104**(5): p. 953-959.
21. Fiorentino, D.F., M.W. Bond, and T.R. Mosmann, *Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones*. J Exp Med, 1989. **170**(6): p. 2081-95.

22. Vieira, P., et al., *Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRF1*. Proc Natl Acad Sci U S A, 1991. **88**(4): p. 1172-6.
23. Hsu, D.H., et al., *Expression of interleukin-10 activity by Epstein-Barr virus protein BCRF1*. Science, 1990. **250**(4982): p. 830-2.
24. Rode, H.J., et al., *The genome of equine herpesvirus type 2 harbors an interleukin 10 (IL10)-like gene*. Virus Genes, 1993. **7**(1): p. 111-6.
25. Fleming, S.B., et al., *A homolog of interleukin-10 is encoded by the poxvirus orf virus*. J Virol, 1997. **71**(6): p. 4857-61.
26. Liu, Y., et al., *The EBV IL-10 homologue is a selective agonist with impaired binding to the IL-10 receptor*. J Immunol, 1997. **158**(2): p. 604-13.
27. Kim, J.M., et al., *Structure of the mouse IL-10 gene and chromosomal localization of the mouse and human genes*. J Immunol, 1992. **148**(11): p. 3618-23.
28. Pestka, S., et al., *Interleukin-10 and related cytokines and receptors*. Annu Rev Immunol, 2004. **22**: p. 929-79.
29. Tone, M., et al., *IL-10 gene expression is controlled by the transcription factors Sp1 and Sp3*. J Immunol, 2000. **165**(1): p. 286-91.
30. Powell, M.J., et al., *Posttranscriptional regulation of IL-10 gene expression through sequences in the 3'-untranslated region*. J Immunol, 2000. **165**(1): p. 292-6.
31. Zdanov, A., C. Schalk-Hihi, and A. Wlodawer, *Crystal structure of human interleukin-10 at 1.6 Å resolution and a model of a complex with its soluble receptor*. Protein Sci, 1996. **5**(10): p. 1955-62.
32. Commins, S., J.W. Steinke, and L. Borish, *The extended IL-10 superfamily: IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28, and IL-29*. J Allergy Clin Immunol, 2008. **121**(5): p. 1108-11.
33. Tan, J.C., et al., *Characterization of interleukin-10 receptors on human and mouse cells*. J Biol Chem, 1993. **268**(28): p. 21053-9.
34. Josephson, K., N.J. Logsdon, and M.R. Walter, *Crystal structure of the IL-10/IL-10R1 complex reveals a shared receptor binding site*. Immunity, 2001. **15**(1): p. 35-46.
35. Usacheva, A., et al., *Two distinct domains within the N-terminal region of Janus kinase 1 interact with cytokine receptors*. J Immunol, 2002. **169**(3): p. 1302-8.
36. Kotenko, S.V., et al., *Other kinases can substitute for Jak2 in signal transduction by interferon-gamma*. J Biol Chem, 1996. **271**(29): p. 17174-82.
37. Finbloom, D.S. and K.D. Winestock, *IL-10 induces the tyrosine phosphorylation of tyk2 and Jak1 and the differential assembly of STAT1 alpha and STAT3 complexes in human T cells and monocytes*. J Immunol, 1995. **155**(3): p. 1079-90.
38. Wehinger, J., et al., *IL-10 induces DNA binding activity of three STAT proteins (Stat1, Stat3, and Stat5) and their distinct combinatorial assembly in the promoters of selected genes*. FEBS Lett, 1996. **394**(3): p. 365-70.
39. Ding, Y., et al., *Suppressor of cytokine signaling 1 inhibits IL-10-mediated immune responses*. J Immunol, 2003. **170**(3): p. 1383-91.
40. Islam, H., et al., *Elevated Interleukin-10 Levels in COVID-19: Potentiation of Pro-Inflammatory Responses or Impaired Anti-Inflammatory Action?* Front Immunol, 2021. **12**: p. 677008.
41. Saraiva, M., P. Vieira, and A. O'Garra, *Biology and therapeutic potential of interleukin-10*. J Exp Med, 2020. **217**(1).
42. Moore, K.W., et al., *Interleukin-10 and the interleukin-10 receptor*. Annu Rev Immunol, 2001. **19**: p. 683-765.
43. de Waal Malefyt, R., et al., *Interleukin 10 (IL-10) and viral IL-10 strongly reduce antigen-specific human T cell proliferation by diminishing the antigen-presenting capacity of monocytes via*

- downregulation of class II major histocompatibility complex expression.* J Exp Med, 1991. **174**(4): p. 915-24.
44. de Waal Malefyt, R., et al., *Interleukin 10(IL-10) inhibits cytokine synthesis by human monocytes: an autoregulatory role of IL-10 produced by monocytes.* J Exp Med, 1991. **174**(5): p. 1209-20.
 45. Hutchins, A.P., D. Diez, and D. Miranda-Saavedra, *The IL-10/STAT3-mediated anti-inflammatory response: recent developments and future challenges.* Brief Funct Genomics, 2013. **12**(6): p. 489-98.
 46. Itoh, K. and S. Hirohata, *The role of IL-10 in human B cell activation, proliferation, and differentiation.* J Immunol, 1995. **154**(9): p. 4341-50.
 47. Sabat, R., et al., *Biology of interleukin-10.* Cytokine Growth Factor Rev, 2010. **21**(5): p. 331-44.
 48. Kuhn, R., et al., *Interleukin-10-deficient mice develop chronic enterocolitis.* Cell, 1993. **75**(2): p. 263-74.
 49. van Montfrans, C., L. Camoglio, and S.J. van Deventer, *Immunotherapy of Crohn's disease.* Mediators Inflamm, 1998. **7**(3): p. 149-52.
 50. Asadullah, K., et al., *IL-10 is a key cytokine in psoriasis. Proof of principle by IL-10 therapy: a new therapeutic approach.* J Clin Invest, 1998. **101**(4): p. 783-94.
 51. Katsikis, P.D., et al., *Immunoregulatory role of interleukin 10 in rheumatoid arthritis.* J Exp Med, 1994. **179**(5): p. 1517-27.
 52. DeBruyne, L.A., et al., *Lipid-mediated gene transfer of viral IL-10 prolongs vascularized cardiac allograft survival by inhibiting donor-specific cellular and humoral immune responses.* Gene Ther, 1998. **5**(8): p. 1079-87.
 53. Rosenblum, I.Y., R.C. Johnson, and T.J. Schmahai, *Preclinical safety evaluation of recombinant human interleukin-10.* Regul Toxicol Pharmacol, 2002. **35**(1): p. 56-71.
 54. Galeazzi, M., et al., *A phase IB clinical trial with Dekavil (F8-IL10), an immunoregulatory 'armed antibody' for the treatment of rheumatoid arthritis, used in combination with methotrexate.* Isr Med Assoc J, 2014. **16**(10): p. 666.
 55. Mumm, J.B., et al., *IL-10 elicits IFN γ -dependent tumor immune surveillance.* Cancer Cell, 2011. **20**(6): p. 781-96.
 56. Emmerich, J., et al., *IL-10 directly activates and expands tumor-resident CD8(+) T cells without de novo infiltration from secondary lymphoid organs.* Cancer Res, 2012. **72**(14): p. 3570-81.
 57. Naing, A., et al., *Safety, Antitumor Activity, and Immune Activation of Pegylated Recombinant Human Interleukin-10 (AM0010) in Patients With Advanced Solid Tumors.* J Clin Oncol, 2016. **34**(29): p. 3562-3569.
 58. Cho, J.L., et al., *Allergic asthma is distinguished by sensitivity of allergen-specific CD4+ T cells and airway structural cells to type 2 inflammation.* Sci Transl Med, 2016. **8**(359): p. 359ra132.
 59. Schulke, S., *Induction of Interleukin-10 Producing Dendritic Cells As a Tool to Suppress Allergen-Specific T Helper 2 Responses.* Front Immunol, 2018. **9**: p. 455.
 60. Rousset, F., et al., *Interleukin 10 is a potent growth and differentiation factor for activated human B lymphocytes.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1890-3.
 61. Rousset, F., et al., *Long-term cultured CD40-activated B lymphocytes differentiate into plasma cells in response to IL-10 but not IL-4.* Int Immunol, 1995. **7**(8): p. 1243-53.
 62. Koenig, K.F., et al., *Serum cytokine profile in patients with active lupus nephritis.* Cytokine, 2012. **60**(2): p. 410-6.
 63. Llorente, L., et al., *Clinical and biologic effects of anti-interleukin-10 monoclonal antibody administration in systemic lupus erythematosus.* Arthritis Rheum, 2000. **43**(8): p. 1790-800.
 64. Yoon, S.I., et al., *Epstein-Barr virus IL-10 engages IL-10R1 by a two-step mechanism leading to altered signaling properties.* J Biol Chem, 2012. **287**(32): p. 26586-95.
 65. Jochum, S., et al., *The EBV immunoevasins vIL-10 and BNLF2a protect newly infected B cells from immune recognition and elimination.* PLoS Pathog, 2012. **8**(5): p. e1002704.

66. Incrocci, R., M. McCormack, and M. Swanson-Mungerson, *Epstein-Barr virus LMP2A increases IL-10 production in mitogen-stimulated primary B-cells and B-cell lymphomas*. J Gen Virol, 2013. **94**(Pt 5): p. 1127-1133.
67. Redford, P.S., P.J. Murray, and A. O'Garra, *The role of IL-10 in immune regulation during M. tuberculosis infection*. Mucosal Immunol, 2011. **4**(3): p. 261-70.
68. Mannino, M.H., et al., *The paradoxical role of IL-10 in immunity and cancer*. Cancer Lett, 2015. **367**(2): p. 103-7.
69. Wang, X., et al., *Targeting IL-10 Family Cytokines for the Treatment of Human Diseases*. Cold Spring Harb Perspect Biol, 2019. **11**(2).
70. Jiang, H., et al., *Subtraction hybridization identifies a novel melanoma differentiation associated gene, mda-7, modulated during human melanoma differentiation, growth and progression*. Oncogene, 1995. **11**(12): p. 2477-86.
71. Jiang, H., et al., *The melanoma differentiation associated gene mda-7 suppresses cancer cell growth*. Proc Natl Acad Sci U S A, 1996. **93**(17): p. 9160-5.
72. Su, Z.Z., et al., *The cancer growth suppressor gene mda-7 selectively induces apoptosis in human breast cancer cells and inhibits tumor growth in nude mice*. Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14400-5.
73. Huang, E.Y., et al., *Genomic structure, chromosomal localization and expression profile of a novel melanoma differentiation associated (mda-7) gene with cancer specific growth suppressing and apoptosis inducing properties*. Oncogene, 2001. **20**(48): p. 7051-7063.
74. Wang, M., et al., *Interleukin 24 (MDA-7/MOB-5) signals through two heterodimeric receptors, IL-22R1/IL-20R2 and IL-20R1/IL-20R2*. J Biol Chem, 2002. **277**(9): p. 7341-7.
75. Soo, C., et al., *Cutaneous rat wounds express c49a, a novel gene with homology to the human melanoma differentiation associated gene, mda-7*. J Cell Biochem, 1999. **74**(1): p. 1-10.
76. Zhang, R., Z. Tan, and P. Liang, *Identification of a Novel Ligand-Receptor Pair Constitutively Activated by ras Oncogenes*. Journal of Biological Chemistry, 2000. **275**(32): p. 24436-24443.
77. Schaefer, G., C. Venkataraman, and U. Schindler, *Cutting edge: FISP (IL-4-induced secreted protein), a novel cytokine-like molecule secreted by Th2 cells*. J Immunol, 2001. **166**(10): p. 5859-63.
78. Sarkar, D., et al., *mda-7 (IL-24): signaling and functional roles*. Biotechniques, 2002. **Suppl**: p. 30-9.
79. Lubkowski, J., et al., *Crystal Structure of the Labile Complex of IL-24 with the Extracellular Domains of IL-22R1 and IL-20R2*. J Immunol, 2018. **201**(7): p. 2082-2093.
80. Dumoutier, L., et al., *Cutting Edge: STAT Activation By IL-19, IL-20 and mda-7 Through IL-20 Receptor Complexes of Two Types*. The Journal of Immunology, 2001. **167**(7): p. 3545-3549.
81. Parrish-Novak, J., et al., *Interleukins 19, 20, and 24 Signal through Two Distinct Receptor Complexes*. Journal of Biological Chemistry, 2002. **277**(49): p. 47517-47523.
82. Wang, M. and P. Liang, *Interleukin-24 and its receptors*. Immunology, 2005. **114**(2): p. 166-170.
83. Croker, B.A., H. Kiu, and S.E. Nicholson, *SOCS regulation of the JAK/STAT signalling pathway*. Semin Cell Dev Biol, 2008. **19**(4): p. 414-22.
84. Smith, S., et al., *Interleukin 24: Signal Transduction Pathways*. Cancers (Basel), 2023. **15**(13).
85. Sauane, M., et al., *Mda-7/IL-24 induces apoptosis of diverse cancer cell lines through JAK/STAT-independent pathways*. J Cell Physiol, 2003. **196**(2): p. 334-45.
86. Chada, S., et al., *Bystander activity of Ad-mda7: human MDA-7 protein kills melanoma cells via an IL-20 receptor-dependent but STAT3-independent mechanism*. Mol Ther, 2004. **10**(6): p. 1085-95.
87. Sauane, M., et al., *Melanoma differentiation associated gene-7/interleukin-24 promotes tumor cell-specific apoptosis through both secretory and nonsecretory pathways*. Cancer Res, 2004. **64**(9): p. 2988-93.

88. Davidson, S., et al., *Protein kinase R is an innate immune sensor of proteotoxic stress via accumulation of cytoplasmic IL-24*. *Sci Immunol*, 2022. **7**(68): p. eabi6763.
89. Do, W., et al., *Sigma 1 Receptor plays a prominent role in IL-24-induced cancer-specific apoptosis*. *Biochem Biophys Res Commun*, 2013. **439**(2): p. 215-20.
90. Sie, C., et al., *IL-24 intrinsically regulates Th17 cell pathogenicity in mice*. *J Exp Med*, 2022. **219**(8).
91. Zhong, Y., X. Zhang, and W. Chong, *Interleukin-24 Immunobiology and Its Roles in Inflammatory Diseases*. *Int J Mol Sci*, 2022. **23**(2).
92. Menezes, M.E., et al., *Role of MDA-7/IL-24 a Multifunction Protein in Human Diseases*. *Adv Cancer Res*, 2018. **138**: p. 143-182.
93. Panneerselvam, J., et al., *IL-24 inhibits lung cancer cell migration and invasion by disrupting the SDF-1/CXCR4 signaling axis*. *PLoS One*, 2015. **10**(3): p. e0122439.
94. Bhutia, S.K., et al., *Targeting breast cancer-initiating/stem cells with melanoma differentiation-associated gene-7/interleukin-24*. *Int J Cancer*, 2013. **133**(11): p. 2726-36.
95. Bhutia, S.K., et al., *mda-7/IL-24 differentially regulates soluble and nuclear clusterin in prostate cancer*. *J Cell Physiol*, 2012. **227**(5): p. 1805-13.
96. Nishikawa, T., et al., *Adenovirus-mediated mda-7 (IL24) gene therapy suppresses angiogenesis and sensitizes NSCLC xenograft tumors to radiation*. *Mol Ther*, 2004. **9**(6): p. 818-28.
97. Yacoub, A., et al., *PERK-dependent regulation of ceramide synthase 6 and thioredoxin play a key role in mda-7/IL-24-induced killing of primary human glioblastoma multiforme cells*. *Cancer Res*, 2010. **70**(3): p. 1120-9.
98. Pradhan, A.K., et al., *mda-7/IL-24 Mediates Cancer Cell-Specific Death via Regulation of miR-221 and the Beclin-1 Axis*. *Cancer Res*, 2017. **77**(4): p. 949-959.
99. Cunningham, C.C., et al., *Clinical and local biological effects of an intratumoral injection of mda-7 (IL24; INGN 241) in patients with advanced carcinoma: a phase I study*. *Mol Ther*, 2005. **11**(1): p. 149-59.
100. Emdad, L., et al., *Recent insights into apoptosis and toxic autophagy: The roles of MDA-7/IL-24, a multidimensional anti-cancer therapeutic*. *Semin Cancer Biol*, 2020. **66**: p. 140-154.
101. Su, Z., et al., *Unique aspects of mda-7/IL-24 antitumor bystander activity: establishing a role for secretion of MDA-7/IL-24 protein by normal cells*. *Oncogene*, 2005. **24**(51): p. 7552-66.
102. Whitaker, E.L., et al., *Splice variants of mda-7/IL-24 differentially affect survival and induce apoptosis in U2OS cells*. *Cytokine*, 2011. **56**(2): p. 272-81.
103. Liu, Z., et al., *Engineering T Cells to Express Tumoricidal MDA-7/IL24 Enhances Cancer Immunotherapy*. *Cancer Res*, 2021. **81**(9): p. 2429-2441.
104. Dash, R., et al., *Novel mechanism of MDA-7/IL-24 cancer-specific apoptosis through SARI induction*. *Cancer Res*, 2014. **74**(2): p. 563-74.
105. Gopalan, B., et al., *Nanoparticle based systemic gene therapy for lung cancer: molecular mechanisms and strategies to suppress nanoparticle-mediated inflammatory response*. *Technol Cancer Res Treat*, 2004. **3**(6): p. 647-57.
106. Poindexter, N.J., et al., *Cytokine induction of interleukin-24 in human peripheral blood mononuclear cells*. *J Leukoc Biol*, 2005. **78**(3): p. 745-52.
107. Kumari, S., et al., *Tumor necrosis factor receptor signaling in keratinocytes triggers interleukin-24-dependent psoriasis-like skin inflammation in mice*. *Immunity*, 2013. **39**(5): p. 899-911.
108. Kragstrup, T.W., et al., *The expression of IL-20 and IL-24 and their shared receptors are increased in rheumatoid arthritis and spondyloarthritis*. *Cytokine*, 2008. **41**(1): p. 16-23.
109. Andoh, A., et al., *Expression of IL-24, an activator of the JAK1/STAT3/SOCS3 cascade, is enhanced in inflammatory bowel disease*. *J Immunol*, 2009. **183**(1): p. 687-95.
110. Kumar, N.P., et al., *Type 2 diabetes - Tuberculosis co-morbidity is associated with diminished circulating levels of IL-20 subfamily of cytokines*. *Tuberculosis (Edinb)*, 2015. **95**(6): p. 707-712.

111. Parrish-Novak, J., et al., *Interleukins 19, 20, and 24 signal through two distinct receptor complexes. Differences in receptor-ligand interactions mediate unique biological functions.* J Biol Chem, 2002. **277**(49): p. 47517-23.
112. Tritsarlis, K., et al., *IL-20 is an arteriogenic cytokine that remodels collateral networks and improves functions of ischemic hind limbs.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15364-9.
113. Rutz, S., X. Wang, and W. Ouyang, *The IL-20 subfamily of cytokines--from host defence to tissue homeostasis.* Nat Rev Immunol, 2014. **14**(12): p. 783-95.
114. Myles, I.A., et al., *Signaling via the IL-20 receptor inhibits cutaneous production of IL-1beta and IL-17A to promote infection with methicillin-resistant Staphylococcus aureus.* Nat Immunol, 2013. **14**(8): p. 804-11.
115. Lee, S.J., et al., *Identification of pro-inflammatory cytokines associated with muscle invasive bladder cancer; the roles of IL-5, IL-20, and IL-28A.* PLoS One, 2012. **7**(9): p. e40267.
116. Cui, X.F., X.G. Cui, and N. Leng, *Overexpression of interleukin-20 receptor subunit beta (IL20RB) correlates with cell proliferation, invasion and migration enhancement and poor prognosis in papillary renal cell carcinoma.* J Toxicol Pathol, 2019. **32**(4): p. 245-251.
117. He, Y., et al., *IL-20RB mediates tumoral response to osteoclastic niches and promotes bone metastasis of lung cancer.* J Clin Invest, 2022. **132**(20).
118. Bach, E.A., et al., *Ligand-induced autoregulation of IFN-gamma receptor beta chain expression in T helper cell subsets.* Science, 1995. **270**(5239): p. 1215-8.
119. Schoenborn, J.R. and C.B. Wilson, *Regulation of interferon-gamma during innate and adaptive immune responses.* Adv Immunol, 2007. **96**: p. 41-101.
120. Lin, F.-C. and H.A. Young, *The talented interferon-gamma.* Advances in Bioscience and Biotechnology, 2013. **04**(07): p. 6-13.
121. Langer, J.A., E.C. Cutrone, and S. Kotenko, *The Class II cytokine receptor (CRF2) family: overview and patterns of receptor-ligand interactions.* Cytokine Growth Factor Rev, 2004. **15**(1): p. 33-48.
122. Jung, V., et al., *Human chromosomes 6 and 21 are required for sensitivity to human interferon gamma.* Proc Natl Acad Sci U S A, 1987. **84**(12): p. 4151-5.
123. Cook, J.R., et al., *Sublocalization of the human interferon-gamma receptor accessory factor gene and characterization of accessory factor activity by yeast artificial chromosomal fragmentation.* J Biol Chem, 1994. **269**(9): p. 7013-8.
124. Hemmi, S., et al., *A novel member of the interferon receptor family complements functionality of the murine interferon gamma receptor in human cells.* Cell, 1994. **76**(5): p. 803-10.
125. Hoffmann, H.H., W.M. Schneider, and C.M. Rice, *Interferons and viruses: an evolutionary arms race of molecular interactions.* Trends Immunol, 2015. **36**(3): p. 124-38.
126. Randal, M. and A.A. Kossiakoff, *The structure and activity of a monomeric interferon-gamma:alpha-chain receptor signaling complex.* Structure, 2001. **9**(2): p. 155-63.
127. Thiel, D.J., et al., *Observation of an unexpected third receptor molecule in the crystal structure of human interferon-gamma receptor complex.* Structure, 2000. **8**(9): p. 927-36.
128. Mendoza, J.L., et al., *Structure of the IFNgamma receptor complex guides design of biased agonists.* Nature, 2019. **567**(7746): p. 56-60.
129. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity.* Nature, 1975. **256**(5517): p. 495-7.
130. Parray, H.A., et al., *Hybridoma technology a versatile method for isolation of monoclonal antibodies, its applicability across species, limitations, advancement and future perspectives.* Int Immunopharmacol, 2020. **85**: p. 106639.
131. Bradbury, A. and A. Pluckthun, *Reproducibility: Standardize antibodies used in research.* Nature, 2015. **518**(7537): p. 27-9.
132. Ulmer, K.M., *Protein engineering.* Science, 1983. **219**(4585): p. 666-71.

133. Poljak, R.J., et al., *Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2,8-Å resolution*. Proc Natl Acad Sci U S A, 1973. **70**(12): p. 3305-10.
134. Poljak, R.J., et al., *The three-dimensional structure of the fab' fragment of a human myeloma immunoglobulin at 2.0-ångstrom resolution*. Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3440-4.
135. Morrison, S.L., et al., *Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains*. Proc Natl Acad Sci U S A, 1984. **81**(21): p. 6851-5.
136. Smith, G.P., *Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface*. Science, 1985. **228**(4705): p. 1315-7.
137. McCafferty, J., et al., *Phage antibodies: filamentous phage displaying antibody variable domains*. Nature, 1990. **348**(6301): p. 552-4.
138. Ahamadi-Fesharaki, R., et al., *Single-Chain Variable Fragment-Based Bispecific Antibodies: Hitting Two Targets with One Sophisticated Arrow*. Mol Ther Oncolytics, 2019. **14**: p. 38-56.
139. Yu, X., et al., *Beyond Antibodies as Binding Partners: The Role of Antibody Mimetics in Bioanalysis*. Annu Rev Anal Chem (Palo Alto Calif), 2017. **10**(1): p. 293-320.
140. Ochoa, M.C., et al., *Antibody-dependent cell cytotoxicity: immunotherapy strategies enhancing effector NK cells*. Immunol Cell Biol, 2017. **95**(4): p. 347-355.
141. Gebauer, M. and A. Skerra, *Engineered Protein Scaffolds as Next-Generation Therapeutics*. Annu Rev Pharmacol Toxicol, 2020. **60**: p. 391-415.
142. Monnier, P.P., R.J. Vigouroux, and N.G. Tassew, *In Vivo Applications of Single Chain Fv (Variable Domain) (scFv) Fragments*. Antibodies, 2013. **2**(2): p. 193-208.
143. Conroy, P.J., et al., *Antibodies: From novel repertoires to defining and refining the structure of biologically important targets*. Methods, 2017. **116**: p. 12-22.
144. Fan, G., et al., *Bispecific antibodies and their applications*. J Hematol Oncol, 2015. **8**: p. 130.
145. Wang, Y., et al., *Nanobody-derived nanobiotechnology tool kits for diverse biomedical and biotechnology applications*. Int J Nanomedicine, 2016. **11**: p. 3287-303.
146. Nygren, P.A. and M. Uhlen, *Scaffolds for engineering novel binding sites in proteins*. Curr Opin Struct Biol, 1997. **7**(4): p. 463-9.
147. Skerra, A., *Engineered protein scaffolds for molecular recognition*. J Mol Recognit, 2000. **13**(4): p. 167-87.
148. Gebauer, M. and A. Skerra, *Engineered protein scaffolds as next-generation antibody therapeutics*. Curr Opin Chem Biol, 2009. **13**(3): p. 245-55.
149. Koide, A., et al., *The fibronectin type III domain as a scaffold for novel binding proteins*. J Mol Biol, 1998. **284**(4): p. 1141-51.
150. Beste, G., et al., *Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold*. Proc Natl Acad Sci U S A, 1999. **96**(5): p. 1898-903.
151. Nord, K., et al., *Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain*. Nat Biotechnol, 1997. **15**(8): p. 772-7.
152. Vazquez-Lombardi, R., et al., *Challenges and opportunities for non-antibody scaffold drugs*. Drug Discov Today, 2015. **20**(10): p. 1271-83.
153. Reichert, J.M., *Antibodies to watch in 2015*. MAbs, 2015. **7**(1): p. 1-8.
154. Bedford, R., et al., *Alternative reagents to antibodies in imaging applications*. Biophys Rev, 2017. **9**(4): p. 299-308.
155. Luo, R., H. Liu, and Z. Cheng, *Protein scaffolds: antibody alternatives for cancer diagnosis and therapy*. RSC Chem Biol, 2022. **3**(7): p. 830-847.
156. Simeon, R. and Z. Chen, *In vitro-engineered non-antibody protein therapeutics*. Protein Cell, 2018. **9**(1): p. 3-14.
157. Azhar, A., et al., *Recent advances in the development of novel protein scaffolds based therapeutics*. Int J Biol Macromol, 2017. **102**: p. 630-641.

158. Packer, M.S. and D.R. Liu, *Methods for the directed evolution of proteins*. Nat Rev Genet, 2015. **16**(7): p. 379-94.
159. Binder, U., et al., *High-throughput sorting of an Anticalin library via EspP-mediated functional display on the Escherichia coli cell surface*. J Mol Biol, 2010. **400**(4): p. 783-802.
160. Gai, S.A. and K.D. Wittrup, *Yeast surface display for protein engineering and characterization*. Curr Opin Struct Biol, 2007. **17**(4): p. 467-73.
161. Sidhu, S.S. and S. Koide, *Phage display for engineering and analyzing protein interaction interfaces*. Curr Opin Struct Biol, 2007. **17**(4): p. 481-7.
162. Lipovsek, D. and A. Pluckthun, *In-vitro protein evolution by ribosome display and mRNA display*. J Immunol Methods, 2004. **290**(1-2): p. 51-67.
163. Galan, A., et al., *Library-based display technologies: where do we stand?* Mol Biosyst, 2016. **12**(8): p. 2342-58.
164. Zahnd, C., P. Amstutz, and A. Pluckthun, *Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target*. Nat Methods, 2007. **4**(3): p. 269-79.
165. Kottke, T., et al., *Photoreceptors Take Charge: Emerging Principles for Light Sensing*. Annu Rev Biophys, 2018. **47**: p. 291-313.
166. Ankenbruck, N., et al., *Optochemical Control of Biological Processes in Cells and Animals*. Angew Chem Int Ed Engl, 2018. **57**(11): p. 2768-2798.
167. Seong, J. and M.Z. Lin, *Optobiochemistry: Genetically Encoded Control of Protein Activity by Light*. Annu Rev Biochem, 2021. **90**: p. 475-501.
168. Deisseroth, K., et al., *Next-generation optical technologies for illuminating genetically targeted brain circuits*. J Neurosci, 2006. **26**(41): p. 10380-6.
169. Kneuttinger, A.C., *A guide to designing photocontrol in proteins: methods, strategies and applications*. Biol Chem, 2022. **403**(5-6): p. 573-613.
170. Courtney, T. and A. Deiters, *Recent advances in the optical control of protein function through genetic code expansion*. Curr Opin Chem Biol, 2018. **46**: p. 99-107.
171. Deiters, A., et al., *A genetically encoded photocaged tyrosine*. Angew Chem Int Ed Engl, 2006. **45**(17): p. 2728-31.
172. Bose, M., et al., *The incorporation of a photoisomerizable amino acid into proteins in E. coli*. J Am Chem Soc, 2006. **128**(2): p. 388-9.
173. Rannversson, H., et al., *Genetically encoded photocrosslinkers locate the high-affinity binding site of antidepressant drugs in the human serotonin transporter*. Nat Commun, 2016. **7**: p. 11261.
174. Furter, R., *Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli*. Protein Sci, 1998. **7**(2): p. 419-26.
175. Wang, L., et al., *Expanding the genetic code of Escherichia coli*. Science, 2001. **292**(5516): p. 498-500.
176. Chin, J.W., *Expanding and reprogramming the genetic code*. Nature, 2017. **550**(7674): p. 53-60.
177. Shandell, M.A., Z. Tan, and V.W. Cornish, *Genetic Code Expansion: A Brief History and Perspective*. Biochemistry, 2021. **60**(46): p. 3455-3469.
178. Kwon, D., *How scientists are hacking the genetic code to give proteins new powers*. Nature, 2023. **618**(7966): p. 874-876.
179. Chemla, Y., et al., *Context effects of genetic code expansion by stop codon suppression*. Curr Opin Chem Biol, 2018. **46**: p. 146-155.
180. Noren, C.J., et al., *A general method for site-specific incorporation of unnatural amino acids into proteins*. Science, 1989. **244**(4901): p. 182-8.
181. Arbely, E., et al., *Photocontrol of tyrosine phosphorylation in mammalian cells via genetic encoding of photocaged tyrosine*. J Am Chem Soc, 2012. **134**(29): p. 11912-5.

182. Luo, J., et al., *Genetic Encoding of Photocaged Tyrosines with Improved Light-Activation Properties for the Optical Control of Protease Function*. *Chembiochem*, 2017. **18**(14): p. 1442-1447.
183. Gautier, A., et al., *Genetically encoded photocontrol of protein localization in mammalian cells*. *J Am Chem Soc*, 2010. **132**(12): p. 4086-8.
184. Wesalo, J.S., et al., *Genetically Encoded Lysine Analogues with Differential Light Sensitivity for Activation of Protein Function*. *ChemPhotoChem*, 2024.
185. Wu, N., et al., *A genetically encoded photocaged amino acid*. *J Am Chem Soc*, 2004. **126**(44): p. 14306-7.
186. Nguyen, D.P., et al., *Genetic encoding of photocaged cysteine allows photoactivation of TEV protease in live mammalian cells*. *J Am Chem Soc*, 2014. **136**(6): p. 2240-3.
187. Lemke, E.A., et al., *Control of protein phosphorylation with a genetically encoded photocaged amino acid*. *Nat Chem Biol*, 2007. **3**(12): p. 769-72.
188. Cheung, J.W., et al., *Genetic Encoding of a Photocaged Histidine for Light-Control of Protein Activity*. *Chembiochem*, 2023. **24**(7): p. e202200721.
189. Zhang, X., et al., *Optical Control of Protein Functions via Genetically Encoded Photocaged Aspartic Acids*. *J Am Chem Soc*, 2023. **145**(35): p. 19218-19224.
190. Yang, X., et al., *Constructing Photoactivatable Protein with Genetically Encoded Photocaged Glutamic Acid*. *Angew Chem Int Ed Engl*, 2023. **62**(40): p. e202308472.
191. Ling, X., et al., *Genetic Encoding of a Photocaged Glutamate for Optical Control of Protein Functions*. *CCS Chemistry*, 2023. **5**(6): p. 1301-1307.
192. Zhu, Y., et al., *Genetically encoded bioorthogonal tryptophan decaging in living cells*. *Nat Chem*, 2024.
193. Liu, C.C. and P.G. Schultz, *Adding new chemistries to the genetic code*. *Annu Rev Biochem*, 2010. **79**: p. 413-44.
194. Guo, Q.R. and Y.J. Cao, *Applications of genetic code expansion technology in eukaryotes*. *Protein Cell*, 2023.
195. Horejsi, Z., et al., *Phosphorylation-dependent PIH1D1 interactions define substrate specificity of the R2TP cochaperone complex*. *Cell Rep*, 2014. **7**(1): p. 19-26.
196. Vorup-Jensen, T., et al., *Structure and allosteric regulation of the alpha X beta 2 integrin I domain*. *Proc Natl Acad Sci U S A*, 2003. **100**(4): p. 1873-8.
197. Kumagai, A., et al., *A bilirubin-inducible fluorescent protein from eel muscle*. *Cell*, 2013. **153**(7): p. 1602-11.
198. Buchko, G.W., et al., *Characterization of two potentially universal turn motifs that shape the repeated five-residues fold--crystal structure of a luminal pentapeptide repeat protein from *Cyanotheca 51142**. *Protein Sci*, 2006. **15**(11): p. 2579-95.
199. Cheung, Y.Y., et al., *Crystal structure of a hyperthermophilic archaeal acylphosphatase from *Pyrococcus horikoshii*--structural insights into enzymatic catalysis, thermostability, and dimerization*. *Biochemistry*, 2005. **44**(12): p. 4601-11.
200. McKay, D.B., et al., *Structure and function of steroid receptor RNA activator protein, the proposed partner of SRA noncoding RNA*. *J Mol Biol*, 2014. **426**(8): p. 1766-1785.
201. Kanagawa, M., et al., *Crystal structures of human secretory proteins ZG16p and ZG16b reveal a Jacalin-related beta-prism fold*. *Biochem Biophys Res Commun*, 2011. **404**(1): p. 201-5.
202. Becker, A.K., et al., *A structure of a collagen VI VWA domain displays N and C termini at opposite sides of the protein*. *Structure*, 2014. **22**(2): p. 199-208.
203. Lam, S.Y., et al., *A rigidifying salt-bridge favors the activity of thermophilic enzyme at high temperatures at the expense of low-temperature activity*. *PLoS Biol*, 2011. **9**(3): p. e1001027.

204. Armstrong, E.H., et al., *Structural basis for ligand regulation of the fatty acid-binding protein 5, peroxisome proliferator-activated receptor beta/delta (FABP5-PPARbeta/delta) signaling pathway*. J Biol Chem, 2014. **289**(21): p. 14941-54.
205. Wang, J., et al., *Time-resolved protein activation by proximal decaging in living systems*. Nature, 2019. **569**(7757): p. 509-513.
206. Wang, J., et al., *CAGE-prox: A Unified Approach for Time-Resolved Protein Activation in Living Systems*. Curr Protoc, 2021. **1**(6): p. e180.
207. Bridge, T., et al., *Site-Specific Encoding of Photoactivity in Antibodies Enables Light-Mediated Antibody-Antigen Binding on Live Cells*. Angew Chem Int Ed Engl, 2019. **58**(50): p. 17986-17993.
208. Jedlitzke, B. and H.D. Mootz, *A Light-Activatable Photocaged Variant of the Ultra-High Affinity ALFA-Tag Nanobody*. Chembiochem, 2022. **23**(12): p. e202200079.
209. Jedlitzke, B., et al., *Photobodies: Light-Activatable Single-Domain Antibody Fragments*. Angew Chem Int Ed Engl, 2020. **59**(4): p. 1506-1510.
210. Joest, E.F., et al., *Light-guided intrabodies for on-demand in situ target recognition in human cells*. Chem Sci, 2021. **12**(16): p. 5787-5795.
211. O'Shea, J.M., et al., *Generation of Photocaged Nanobodies for Intracellular Applications in an Animal Using Genetic Code Expansion and Computationally Guided Protein Engineering*. Chembiochem, 2022. **23**(16): p. e202200321.
212. Bridge, T., et al., *Site-specific encoding of photoactivity and photoreactivity into antibody fragments*. Nat Chem Biol, 2023. **19**(6): p. 740-749.
213. Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces*. J Mol Biol, 1998. **280**(1): p. 1-9.
214. Salomon, R., et al., *Bispecific antibodies increase the therapeutic window of CD40 agonists through selective dendritic cell targeting*. Nat Cancer, 2022. **3**(3): p. 287-302.
215. Driscoll, C.L., A.H. Keeble, and M. Howarth, 2024.
216. Bloom, J.D. and F.H. Arnold, *In the light of directed evolution: pathways of adaptive protein evolution*. Proc Natl Acad Sci U S A, 2009. **106 Suppl 1**(Suppl 1): p. 9995-10000.
217. Zahradnik, J., et al., *A Protein-Engineered, Enhanced Yeast Display Platform for Rapid Evolution of Challenging Targets*. ACS Synth Biol, 2021. **10**(12): p. 3445-3460.
218. Bandari, A.K., et al., *A Novel Splice Site Mutation in IFNGR2 in Patients With Primary Immunodeficiency Exhibiting Susceptibility to Mycobacterial Diseases*. Frontiers in Immunology, 2019. **10**.
219. Chaudhari, A.S., et al., *Genetically encoded non-canonical amino acids reveal asynchronous dark reversion of chromophore, backbone, and side-chains in EL222*. Protein Sci, 2023. **32**(4): p. e4590.
220. Natarajan, C., et al., *Epistasis among adaptive mutations in deer mouse hemoglobin*. Science, 2013. **340**(6138): p. 1324-7.
221. Reetz, M.T., *The importance of additive and non-additive mutational effects in protein engineering*. Angew Chem Int Ed Engl, 2013. **52**(10): p. 2658-66.
222. Mock, J., C. Pellegrino, and D. Neri, *A universal reporter cell line for bioactivity evaluation of engineered cytokine products*. Sci Rep, 2020. **10**(1): p. 3234.
223. Cho, O., et al., *Chelerythrine, a novel small molecule targeting IL-2, inhibits melanoma progression by blocking the interaction between IL-2 and its receptor*. Life Sci, 2023. **320**: p. 121559.
224. Xue, C., et al., *Evolving cognition of the JAK-STAT signaling pathway: autoimmune disorders and cancer*. Signal Transduct Target Ther, 2023. **8**(1): p. 204.
225. Luo, J., et al., *Reversible and Tunable Photoswitching of Protein Function through Genetic Encoding of Azobenzene Amino Acids in Mammalian Cells*. Chembiochem, 2018. **19**(20): p. 2178-2185.

226. Strugach, D.S., D. Hadar, and M. Amiram, *Robust Photocontrol of Elastin-like Polypeptide Phase Transition with a Genetically Encoded Arylazopyrazole*. ACS Synth Biol, 2023. **12**(10): p. 2802-2811.



Article

Protein Binder (ProBi) as a New Class of Structurally Robust Non-Antibody Protein Scaffold for Directed Evolution

Puong Ngoc Pham, Maroš Hulíčiak, Lada Biedermannová, Jiří Černý , Tatsiana Charnavets, Gustavo Fuertes , Štěpán Herynek, Lucie Kolářová , Petr Kolenko , Jiří Pavlíček, Jiří Zahradník, Pavel Mikulecky * and Bohdan Schneider *

Institute of Biotechnology of the Czech Academy of Sciences, BIOCEV, CZ-25250 Vestec, Czech Republic; phuong@ibt.cas.cz (P.N.P.); maros.huliciak@ibt.cas.cz (M.H.); Lada.Biedermannova@ibt.cas.cz (L.B.); jiri.cerny@ibt.cas.cz (J.Č.); tatsiana.charnavets@ibt.cas.cz (T.C.); gustavo.fuertes@ibt.cas.cz (G.F.); Stepan.Herynek@ibt.cas.cz (Š.H.); kolarovalucie007@gmail.com (L.K.); kolenko@ibt.cas.cz (P.K.); jiri.pavlicek@ibt.cas.cz (J.P.); jiri.zahradnik@ibt.cas.cz (J.Z.)

* Correspondence: mikulecky@outlook.com (P.M.); bohdan.schneider@gmail.com (B.S.)

Abstract: Engineered small non-antibody protein scaffolds are a promising alternative to antibodies and are especially attractive for use in protein therapeutics and diagnostics. The advantages include smaller size and a more robust, single-domain structural framework with a defined binding surface amenable to mutation. This calls for a more systematic approach in designing new scaffolds suitable for use in one or more methods of directed evolution. We hereby describe a process based on an analysis of protein structures from the Protein Data Bank and their experimental examination. The candidate protein scaffolds were subjected to a thorough screening including computational evaluation of the mutability, and experimental determination of their expression yield in *E. coli*, solubility, and thermostability. In the next step, we examined several variants of the candidate scaffolds including their wild types and alanine mutants. We proved the applicability of this systematic procedure by selecting a monomeric single-domain human protein with a fold different from previously known scaffolds. The newly developed scaffold, called ProBi (Protein Binder), contains two independently mutable surface patches. We demonstrated its functionality by training it as a binder against human interleukin-10, a medically important cytokine. The procedure yielded scaffold-related variants with nanomolar affinity.

Keywords: directed evolution; protein scaffold; protein engineering; computational saturation; ribosome display; interleukin-10



Citation: Pham, P.N.; Hulíčiak, M.; Biedermannová, L.; Černý, J.; Charnavets, T.; Fuertes, G.; Herynek, Š.; Kolářová, L.; Kolenko, P.; Pavlíček, J.; et al. Protein Binder (ProBi) as a New Class of Structurally Robust Non-Antibody Protein Scaffold for Directed Evolution. *Viruses* **2021**, *13*, 190. <https://doi.org/10.3390/v13020190>

Academic Editor: Petri Susi
Received: 19 December 2020
Accepted: 23 January 2021
Published: 27 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many biological functions depend on specific protein-protein interactions. Protein engineering offers the possibility to tune these interactions by developing de novo binding partners [1–3] or by mutating the interaction partners using computational design [4,5]. A powerful tool of protein engineering is to generate protein binders by the in vitro directed selection techniques [6–8] or to use evolution-based approaches to increase the stability of recombinant proteins [9–11].

A prototype of a binding protein is an antibody, a highly selective and adaptive molecule capable of binding to a huge spectrum of partners. The antibody-based binders have been indispensable not only in research experiments but also in clinical trials and tens of them are biologicals approved by Federal Drug Administration. However, several suboptimal properties of these molecules such as their large size (~150 kDa), cross-reactivity, and necessity of animal immunization during the preparation, motivated the development of binders with alternative structures [12,13]. These novel artificial high-affinity binders are called “small non-antibody protein scaffolds” [14–16]. The advantage of scaffolds over antibodies is their typically higher stability and lower cost of production. The most widely

used scaffolds are designed Ankyrin repeats (dARPs), lipocalin domain (anticalins), and Ig-binding domain of Staphylococcal Protein A (affibodies) [17–19]. Some of the small protein scaffolds were already proven to be useful in a wide range of applications, from academic research to clinical imaging, diagnostics, or therapeutics [20–23], and others are even in clinical trials [24,25]. In our previous work, we used the albumin-binding domain (ABD) of Streptococcal Protein G as a scaffold protein. We engineered its binding against human IFN γ [26] and interleukin-23 receptor [27] using computational design and ribosome display.

The exploration of non-antibody protein scaffolds is still a young field compared to the research of antibodies and the full potential of scaffolds has not yet been uncovered. Therefore, we are convinced that there is ample space to expand the current portfolio of protein scaffolds so that they can complement antibodies and interact with new target molecules. Therefore, a systematic method to streamline the development of new scaffolds with desired properties would be beneficial but relatively few attempts can be found in the literature [18,28,29]. In general, a good small protein scaffold should display a high stability, together with flexibility of binding. The desired properties thus include tolerance to mutations during diversification, thermostability, *in vivo* integrity, and non-immunogenicity. Moreover, the variants of the protein scaffold must be efficiently evolvable to achieve high affinity and specificity towards target proteins. Therefore, the properties of the potential small protein scaffold must be thoroughly evaluated.

In this work, we present a systematic stepwise procedure for the selection of new and structurally robust protein scaffolds. The procedure starts with a wide search of suitable structures in the Protein Data Bank (PDB) [30], followed by a series of computational and experimental screens to select the most robust candidates as novel, small protein scaffolds. We demonstrate the applicability of our method by selecting a protein scaffold called ProBi (Protein Binder) that fulfills all the criteria of stability and mutability, and comprises not one, but two independently mutable surface patches.

We tested the applicability of the ProBi scaffold by generating its variants with high affinity to a target protein using *in vitro* directed evolution techniques. As the target, we selected a medically important protein, interleukin-10 (IL-10), a member of the cytokine superfamily. IL-10 is an immune repressor [31], with numerous investigations suggesting its major impact in inflammatory, malignant, and autoimmune diseases. IL-10 overexpression was found in certain tumors and is considered to promote tumor development [32]. Also, some viral homologs can bind to the IL-10 receptor 1 [33] and mimic the immunosuppressive effects of IL-10. Thus, proteins binding IL-10 and interfering with its signaling have a great potential for diagnostic and therapeutic purposes.

To generate ProBi-based binders to IL-10, we created a DNA library by randomization of 10 amino acids on one of the ProBi surface patches and used the ribosome display technique to screen for ProBi variants with the highest affinity. We successfully generated binders of the newly developed ProBi protein scaffold with affinities of ~10 and ~200 nM against IL-10.

2. Materials and Methods

2.1. Selection of Suitable Protein Structures from the Protein Data Bank

The Protein Data Bank (PDB) [30] was searched for potential protein scaffolds using the following criteria: (1) small size (molecular weight in range of 10–25 kDa); (2) monomeric protein with a reasonably high-resolution X-ray structure (<3.0 Å); and (3) produced in Expression Organism of *E. coli* for ease of expression. Proteins matching these initial search criteria were then manually curated to meet additional rules: structure (protein fold class) (1) should differ from the previously published protein scaffolds; (2) should differ from the other selected structures in the set; (3) should contain a low number of cysteine residues and disulfide bonds, (4) should not be toxic or immunogenic, (5) should not contain any ligand or cofactor, and (6) should be soluble and easy to purify (according to the source literature).

2.2. In Silico Identification of Mutable Surface Patches on Protein Scaffolds Candidates

(a) Multiple sequence alignment. For each of the scaffold candidate proteins, a similarity search was performed for the related amino acid sequences from different species using the UniProt BLAST service (www.uniprot.org/blast) with default parameters (Target database = UniProtKB, E-Threshold = 10, Matrix = blosum62, Filtering = None, Gapped = yes, Hits = 250). The identical sequences were subsequently removed using a script and the remaining ones were aligned using Clustal W [34] in the UGENE program [35]. The conservation of each sequence position was calculated.

(b) In silico saturation mutagenesis. The mutability of the surface amino acid residues was evaluated by in silico mutation scanning using the FoldX program [36] by applying the “positionscan” FoldX Keywords: In all proteins, each position was substituted by all of the 20 standard amino acid residues, and the corresponding free energy differences ($\Delta\Delta G$) between the mutant and the wild-type (WT) structure were calculated; the calculations included self-mutations leading to $\Delta\Delta G = 0$. Positions in which most of the mutations were stabilizing ($\Delta\Delta G < 0$), or only slightly destabilizing ($\Delta\Delta G < 0.5$ kcal/mol), were identified and the mutability score for each position was calculated as a percentage of mutations fulfilling these criteria.

(c) Sequence and structural evaluation. The mutable residues, identified by multiple sequence alignment and in silico saturation mutagenesis, were located on the 3D structure of each protein scaffold candidate. We then visually searched for continuous surface regions consisting of mutable residues. The regions where the mutable residues were close together in both sequential and structural space we termed the mutable surface patch. Ideally, the patch should contain 10–12 residues to fulfill the theoretical complexity suitable for ribosome display. If the identified patch contained less than 10–12 residues with the highest mutability score, we completed the patch by a small number of neighboring surface residues with slightly lower mutability scores.

2.3. Production of Proteins

(a) Cloning of recombinant proteins. The genes coding the selected five protein scaffold candidates—4PSF, 1N3Y, 4I3B, 1W2I, and 2F3L, were ordered in the form of DNA strings (ThermoFisher Scientific, Waltham, MA, USA) with codons optimized for expression in *Escherichia coli* but without a stop codon. The DNA strings were cloned into the pET-26b(+) vector using NdeI and XhoI restriction enzymes in frame with C-terminal His-tag. Also, the DNA strings were used as templates for polymerase chain reaction (PCR) to add the N-terminal His-tag and stop codon. These N-terminally tagged proteins were cloned into the pET-26b(+) vector using NdeI and XhoI restriction enzymes. Final constructs were labeled as pET26b-4PSF-WT, etc.

The genes coding the variants of three protein scaffold candidates—4PSF, 1N3Y, and 4I3B, with all mutable residues changed to alanine amino acid, referred to as allAla mutants, were ordered in the form of DNA strings with N-terminal His-tag and C-terminal stop codon. The DNA strings were digested by NdeI and XhoI restriction enzymes and cloned into the pET-26b(+) vector. Final constructs were labeled as pET26b-4PSF-allAla, etc.

The ProBi-WT protein and its variants were amplified by the PCR method using the ProBi-cloning-for and ProBi-cloning-rev primers (Table S1). The created fragments were cleaved by NcoI and BamHI restriction enzymes and cloned into the pRDVsm or pETsm vectors (Figure S1).

The genes encoding the interleukin-10 (residues 19–179 of UniProt entry P22301) and interleukin-29 (residues 20–200 of UniProt entry Q8IU54) were ordered in the form of a DNA string with codons optimized for expression in *Drosophila melanogaster*. The DNA was cleaved by BglII and XhoI restriction enzymes and cloned into a pMTH vector (a modified version of pMT-BiP-V5-His_A vector where recognition site for the AgeI restriction enzyme is exchanged for XhoI enzyme). The final construct contained IL-10 or IL-29 protein with N-terminal BiP signal peptide and C-terminal His-tag. Final constructs were labeled as pMTH-IL10cd and pMTH-IL29cd.

(b) Expression and purification of recombinant proteins. The recombinant protein scaffold candidates were produced in *E. coli* BL21(DE3) strain. In addition, the expression of 1W3I scaffold was tested in *E. coli* C41(DE3) and C43(DE3) strains. The bacteria were cultivated in LB medium (Sigma-Aldrich, St. Louis, MO, USA) at 37 °C until OD₆₀₀ = 0.6, followed by the addition of 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for induction of expression. Then the cultivation continued for 4 h at 37 and 30 °C or overnight at 25 and 16 °C. The cells were harvested by centrifugation (5000 × *g*, 15 min, 4 °C) and stored at −20 °C. The cells were resuspended in LH buffer (20 mM Na-Phosphate buffer, pH 7.3, 100 mM NaCl), and disrupted by ultrasound. The soluble fraction was separated by centrifugation (40,000 × *g*, 20 min, 4 °C) and passed over the Ni-NTA agarose beads equilibrated in LH buffer. The beads were washed by LH buffer supplemented by 10 mM imidazole and proteins were eluted by LH buffer containing 200 mM imidazole.

The recombinant ProBi variants were produced in *E. coli* BL21-Gold(DE3) in LB medium (Sigma-Aldrich) at 37 °C until OD₆₀₀ = 0.6, followed by the addition of 1 mM IPTG for induction of expression. Then cultivation continued overnight at 18 °C. The cells were harvested by centrifugation (5000 × *g*, 15 min, 4 °C) and stored at −20 °C. The cells were resuspended in LS buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA), and disrupted by ultrasound. The soluble fraction was separated by centrifugation (40,000 × *g*, 20 min, 4 °C) and passed over the StrepTactin XT resin equilibrated in LS buffer. The resin was washed by LS buffer and proteins were eluted by BXT buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM biotin).

The constructs pMTH-IL10cd and pMTH-IL29cd for expression of recombinant IL-10 and IL-29, respectively, in insect S2 cells were produced and purified similarly as IFNgR2 in the previous report [37]. Briefly, S2 cells were cultivated in Insect-XPRESSTM Protein-free Insect Cell Medium (Lonza) and protein expression was induced by the addition of CuSO₄. IL-10 and IL-29 proteins were secreted into the medium and stored at −80 °C.

All proteins were further purified to homogeneity by size exclusion chromatography at 16 °C using Superdex 75 10/300 Increase, Superdex 200 10/300 Increase, or Superdex 75 16/600 column (GE Healthcare, Chicago, IL, USA). The column was equilibrated in phosphate-buffered saline (PBS) pH 7.4 (for protein scaffold proteins and all Ala mutants), 20 mM Tris, pH 7.5, 100 NaCl (for preliminary crystallization trials), or in 50 mM Tris, pH 8.0, 300 mM NaCl (for ProBi variants and IL-10 protein). Samples were analyzed by 15% sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).

2.4. Biophysical Characterization of Recombinant Proteins

(a) Oligomerization and melting temperature of protein scaffold candidates and allAla mutants. The oligomeric state of proteins and their monodispersity was measured by dynamic light scattering (DLS) using a ZetaSizer Nano ZS90 (Malvern Panalytical, Malvern, UK) instrument and quartz cuvette ZEN2112. The secondary structure and folding of proteins were analyzed by circular dichroism (CD) spectra using a Chirascan Plus spectrometer (Applied Photophysics, Surrey, UK) in steps of 1 mm over the wavelength range 200–260 nm. Samples were diluted with PBS buffer, pH 7.4, to a concentration of 0.2 mg/mL. The individual spectra were measured in a 0.01 cm path-length quartz cell at a temperature of 25 °C. The CD signal was recorded as the ellipticity, and the resulting spectra were buffer-subtracted. The melting temperatures of proteins were evaluated using either the CD spectrometer or nanoDSF technology. The CD melting measurements were performed by using the 0.1 cm path-length quartz cell. The samples were heated from 20 to 85 °C, and the sample absorption was recorded at 280 nm in 1 °C increments at a rate of 0.5 °C/min. The nanoDSF technology was implemented in the Prometheus NT.48 instrument (NanoTemper, München, Germany). The samples were loaded into standard capillaries and heated from 20 to 95 °C at a rate of 1 °C/min. The melting temperatures (*T_m*) were estimated from the first derivative of the melting curves.

(b) Testing protein suitability for affinity measurements. The applicability of protein scaffold candidates and allAla mutants for affinity measurements using the microscale

thermophoresis (MST) and surface plasmon resonance (SPR) was tested. For MST, proteins were labeled using the Monolith Protein Labeling Kit RED-NHS (Amine Reactive) according to the manufacturer's protocol. The samples were loaded into three types of capillary (standard, hydrophilic, and hydrophobic) and inserted into the Monolith NT.115 instrument (NanoTemper). The capillary scan was performed immediately, and then samples were incubated in the instrument for 1 h and re-measured. An alternative approach was implemented for SPR. The GLC sensor chip was inserted into the ProteOn XPR36 instrument, washed but not activated. The samples in different channels were passed over the un-activated chip to estimate the non-specific binding of proteins to the surface of the chip itself. The running buffer (PBS, pH 7.4, 0.005% Tween-20) was used in the control channel.

2.5. Crystallization and Structure Determination

The crystallization solution included J61 at a concentration of 10 mg/mL in a buffer containing 100 mM NaCl, 20 mM Tris, pH 8.0. Crystals were prepared using the hanging-drop vapor-diffusion method with the reservoir solution containing 5% (*v/v*) glycerol and 4 M sodium formate. Crystals required cryo-protection with 20% glycerol before flash freezing in liquid nitrogen and the diffraction experiment.

The final diffraction data were collected at the Helmholtz-Zentrum Berlin (Bessy II) on a beamline 14.1 at 100 K. The diffraction images were processed using the XDS program [38] and scaled using the AIMLESS program [39] from the CCP4 program package [40]. The phase problem was solved by molecular replacement using PHASER [41] using monomeric 4PSF structure [42] as a search model. The structure was refined using the REFMAC5 program [43] using 95% of reflections as a working set and 5% or reflections as a test set. Manual corrections to the model were done using the graphic COOT program [44]. The last run of several cycles of refinement was performed using all reflections and anisotropic ADPs. The structure quality was checked using the validation tools implemented in CCP4, MOLPROBITY [45], and the PDB. Final data processing statistics and structure refinement statistics are shown in Table S2.

2.6. Construction of DNA Library of ProBi Scaffold (PatchC)

The earlier design of the DNA library [46] was adopted with small changes regarding restriction enzyme recognition sites and N- and C-terminal tags. Our ribosome display DNA construct contained a T7 promoter, 5' stem-loop, ribosome binding site, N-terminal Strep-tag, open reading frame encoding the ProBi scaffold library, and C-terminal c-Myc-tag fused to TolA-spacer without stop codon. The library of ProBi scaffold (PatchC) was prepared by the GENEWIZ company using NNK codons technology. We selected ten residues to be randomized. The initial synthetic DNA library was amplified by PCR and then used in the first selection round of ribosome display.

2.7. Selection of Novel Binders by In Vitro Ribosome Display

The previous protocols from Pluckthun's laboratory regarding the in vitro selection by ribosome display [46,47] were adopted and adjusted. All the following steps were first performed with the ProBi-WT scaffold to establish and verify the protocol.

(a) Preparation of the mRNA-ribosome-protein (MRP) complex. The DNA library was amplified by PCR (supplemented by 6% DMSO) with primers T7b and TolAk (Table S1) using Q5 polymerase (New England Biolabs, Ipswich, MA, USA) with an annealing temperature of 66 °C. As a template, either the starting DNA library in the first selection round was used or the sorted DNA library cloned into the pRDVsm vector in the next rounds. The resulted amplicon contained a T7 promoter, 5' stem-loop, ribosome binding site (RBS), start codon, N-terminal Strep-tag II, scaffold library, C-terminal c-Myc-tag, tolA spacer, and 3' stem-loop without stop codon. The amplicon was purified from 1% agarose gel and used in the following cell-free protein synthesis step. We utilized RTS 100 *E. coli*

HY Kit (Biotectrabbit) according to manufacturer recommendations to create the MRP complex composed of mRNA, ribosome, and protein scaffold variants.

(b) Ribosome display selection rounds. The Nunc PolySorp plate (Invitrogen, Carlsbad, CA, USA) was coated with anti-His antibody (His-tag Monoclonal Antibody (HIS.H8) by Invitrogen) diluted to a final concentration of 25 µg/mL in bicarbonate coating buffer (pH 9.6). All following steps were undertaken in the TBS buffer, pH 7.4. Blocking of plate and incubation of scaffold variants with proteins were undertaken either by an initial preselection process or an optimized preselection process described below.

(c) Initial preselection process. The plate was blocked by 3% bovine serum albumin (BSA) during all rounds. The MRP complex supplemented by 0.5% BSA and heparin (200 mg/mL) was added to the Preselection well that was coated with a mixture of 3% BSA, interferon-gamma (IFN γ , 25 µg/mL), and the 4PSF wild type protein (4PSF-WT, 25 µg/mL). The reaction was incubated for 1 h at 4 °C.

(d) Optimized preselection process. The plate was blocked by either 3% BSA (used in the first and third round) or by 3% dry skimmed cow milk (used in the second and fourth round). The MRP complex supplemented by 0.5% BSA and heparin (200 mg/mL) was added to the Preselection well that was coated with a mixture of 3% BSA, IFN γ (25 µg/mL), IL-29 (25 µg/mL), and 4PSF-WT (25 µg/mL). The reaction was incubated for 1 h at 4 °C.

(e) Selection process. The preselected MRP complex was transferred into the Selection well that contained IL-10 (25 µg/mL) as a target molecule and incubated again for 1 h at 4 °C. In each round, the well was washed by different cycles of washes and increasing concentration of Pluronic F-127 in TBS buffer. The conditions were as follows: First round–5 wash cycles with 0.10%, second round–10 wash cycles with 0.15%, third and fourth round–10 wash cycles with 0.20%. The library complex was incubated with an EB buffer (50 mM Tris-Acetate, 150 mM NaCl, 50 mM EDTA, pH adjusted to 7.5) containing *S. cerevisiae* RNA (1 mg/mL) and heparin (200 mg/mL) to release the mRNA. The mRNA was purified by the High Pure RNA Isolation Kit (Roche, Basel, Switzerland) according to the manufacturer's instructions and used in the GoScript Reverse Transcription System (Promega, Dane County, WI, USA) with ProBi-cloning-rev primer according to the manufacturer's instruction. The created cDNA was amplified by PCR (supplemented by 6% DMSO) with primers ProBi-cloning-for and ProBi-cloning-rev using Q5 polymerase (New England Biolabs). The resulted amplicon was cloned into the pRDVsm vector using NcoI and BamHI restriction enzymes and T4 ligase (New England Biolabs). The pRDVsm-ProBi was used in the following preselection and selection process. After the fourth round, the amplicon was cloned into the pETsm vector and transformed into BL21(DE3)-Gold (Agilent, Santa Clara, CA, USA) for affinity estimation by enzyme-linked immunosorbent assay (ELISA) and characterization.

2.8. Characterization of Novel Binders

(a) ELISA assay. The Nunc PolySorp plate (Invitrogen) was coated with anti-His antibody (His-tag Monoclonal Antibody (HIS.H8) by Invitrogen) diluted to a final concentration of 5 µg/mL in bicarbonate coating buffer (pH 9.6). All following steps were done in the PBS-P buffer (PBS buffer, pH 7.4, 0.1% Pluronic F-127). For the first evaluation of the binding of 190 variants, the wells were blocked by 1% BSA, then IL-10 (10 µg/mL) was attached as a target molecule, and samples were applied in the form of cell lysates. For initial specificity mapping of 47 variants, the wells were blocked by 3% dry skim milk, 3% BSA, or 1% BSA. The wells with 1% BSA were further incubated with either IL-10 or IL-29 (10 µg/mL), and samples were applied in the form of cell lysates. For affinity estimation of 10 selected variants, the wells were blocked by 1% BSA and then incubated with IL-10 (10 µg/mL) as a target molecule. Then the samples in the form of proteins purified on StrepTactinXT were applied. Finally, the binders were detected by horseradish peroxidase (HRP)-conjugated antibody against the C-terminal c-Myc-tag and TMB-2 as substrate.

(b) Inhibitory assay. Murine RAW264.7 cells were cultured in DMEM medium supplemented with 10% FBS. The mixture containing IL-10 (final concentration of 10 ng/mL)

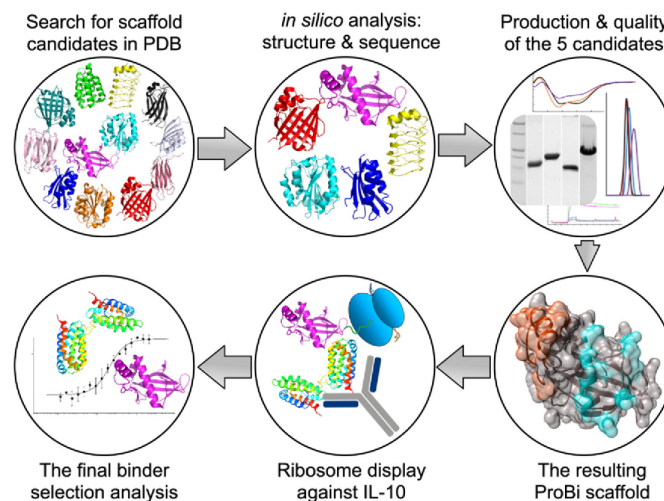
and ProBi variants (concentrations ranging from 160 ng to 100 µg/mL) was added to the cells and incubated for 30 min. The plates with cells were placed on ice, medium discarded, and washed with a cold PBS buffer. The cells were disrupted by a cold RIPA buffer supplemented by protease and phosphatase inhibitors. An amount of 20 µg of total protein was resolved using 10% Tris-Glycine SDS-PAGE gel and transferred to nitrocellulose membrane using Trans-Blot Turbo Transfer System according to the manufacturer's instructions. Membranes were blocked by 5% BSA in TBST buffer and incubated with antibodies against Y705-STAT3 (dilution 1/700) and alpha-tubulin (dilution 1/5000) diluted in 5% BSA in TBST buffer. Blots were washed in TBST buffer and incubated with appropriate HRP-conjugated secondary antibodies diluted in 1% BSA in TBST buffer. Membranes were treated with ECL solution (Merck Millipore, Burlington, MA, USA) and visualized on the Azure c600 instrument.

(c) Affinity measurements. The binding affinity in a solution was measured using the Monolith NT.115 instrument to monitor microscale thermophoresis (MST). The composition of the Assay buffer was 50 mM Tris, pH 8.0, 300 mM NaCl, and 0.1% Pluronic F-127. The IL-10 with C-terminal His-tag was labeled in assay buffer using the Monolith His-tag Labeling Kit RED-tris-NTA kit according to the manufacturer's instructions. The labeled IL-10 was titrated by ProBi variants diluted in an assay buffer and incubated for 10 min at room temperature. The reaction mix was loaded into standard capillaries and MST was monitored using Medium MST power and 60% of LED power. The analysis of interaction was done in MO.Affinity Analysis v2.3 software.

(d) Circular dichroism. The folding of new ProBi variants was measured by Circular dichroism spectra the same way as done with allAla mutants of protein scaffold candidates but with a final concentration of 0.1 mg/mL.

3. Results and Discussion

This work aimed to formulate and test a general procedure for selecting “small engineered non-antibody protein scaffolds” [15,16] suitable for generating artificial recombinant high-affinity binders by methods of targeted *in vitro* evolution. The systematic procedure for the development and testing of a new stable small protein scaffold presented in this work consists of two main parts comprising the steps graphically represented in Scheme 1 and summarized below:



Scheme 1. The procedure for the development of a new stable small protein scaffold.

(I) Development of a novel scaffold:

- (i) Selection of suitable protein candidates from the Protein Data Bank.
- (ii) In silico identification of mutable patches on the surfaces of the scaffold candidates.
- (iii) Characterization of soluble wild-type variants of the scaffold candidates.
- (iv) Testing the stability of the potential scaffolds.
- (v) Testing the suitability of the scaffold candidate(s) for in vitro evolution.
- (vi) Construction of the final ProBi scaffold.
- (II) Application of the ProBi scaffold on the IL-10 target as a model system:
 - (i) Selection of novel ProBi-based binders by ribosome display.
 - (ii) Characterization of the novel ProBi-based binders.
 - (iii) Further development of scaffold selection and improvement of ProBi-based binders.

3.1. Development of a Novel Scaffold

3.1.1. Selection of Suitable Protein Structures from the Protein Data Bank

The first step of the procedure was to search for appropriate structures within the Protein Data Bank (PDB) database [30] performed by the members of the laboratory. The initial search was based on visual inspection of several hundred structures of small (less than 25 kDa) monomeric proteins solved by X-ray crystallography. The search results were limited to proteins known to be expressed in *Escherichia coli*. We finally selected about 100 structures that were further evaluated based on the literature data. Our selected candidates should (i) be new protein scaffolds, (ii) be based on different folds, (iii) preferably contain no or low number of cysteines, (iv) not to be toxic, (v) not to need any cofactor to maintain the structure, and (vi) be easy to purify.

We chose 12 proteins (Table 1) for closer structural examination and a more in-depth literature review. During this step, we excluded 2W4P [48] and 4LKT [49] because of structural similarities with other selected non-human candidates, 1W2I [50] and 4I3B [51], respectively. Furthermore, we eliminated 4MJJ and 4JOX because the manuscripts were not published at that time and we could not check their expression and purification protocols. To the best of our knowledge, none of the eight finally selected proteins had previously been used as a scaffold for directed evolution.

3.1.2. In Silico Identification of Mutable Surface Patches on the Protein Scaffold Candidates

On the surface of the eight candidate scaffolds, we looked for continuous patches consisting of residues that may be mutated without destabilizing the overall structure. To identify such mutable residues, we combined the analysis of protein sequence conservation with in silico saturation mutagenesis. We did not limit our search only for loop segments [28] because protein scaffolds could be mutated also at flat surfaces or combinations of loops and helices [18].

(a) Multiple sequence alignment. Sequences of protein homologs from different organisms for the eight protein scaffold candidates were aligned and the conservation of each amino acid position was calculated.

(b) In silico saturation mutagenesis. To evaluate the mutability of the residues, we performed in silico mutation scanning using the FoldX program [36]. Every amino acid position in each of the eight candidates was substituted by all of the 20 standard amino acid residues, and the corresponding free energy differences ($\Delta\Delta G$) between the mutant and the wild-type structure were calculated. Thus, each position in the protein sequence was characterized by 20 $\Delta\Delta G$ values. We identified spots in which most of the mutations were stabilizing ($\Delta\Delta G < 0$), or only slightly destabilizing ($0 < \Delta\Delta G < 0.5$ kcal/mol). We calculated the mutability score for each position as a percentage of mutations fulfilling these criteria.

(c) Sequence and structural evaluation. For the selection of mutable residues, we combined the evolutionary conservation and mutability scores. Positions with identity conservation $< 90\%$ and FoldX mutability score $> 50\%$ were considered mutable. Next, we mapped the mutable residues onto the protein 3D structures. We visually identified mutable positions and selected those constituting compact patches on the protein surface.

We preferred those patches consisting of residues that were not only structurally, but also sequentially close together, for ease of the subsequent library construction. To achieve a high variability of the designed DNA library, we selected 10–12 mutable residues, the numbers typically used in ribosome display experiments [57,58]. In cases where the patches contained less than 8–10 predicted mutable residues, we chose additional neighboring solvent-accessible residues to bring the total number of residues to 10.

Table 1. List of protein scaffold candidates that were selected for closer in silico structural analysis and literature review. The first five candidates were chosen for further experimental characterization.

PDB Code	UniProt	Protein Name	Source Organism	Size (kDa)	Size (aa)	Number of Cysteines	Reference
4PSF	Q9NWS0	PIH1D1 N-terminal domain	<i>Homo sapiens</i>	15	138	3	[42]
1N3Y	P20702	Alpha-X beta2 integrin I domain	<i>Homo sapiens</i>	22	198	0	[52]
4I3B	P0DM59	Fluorescent protein UnaG wild type	<i>Anguilla japonica</i>	15	139	0	[51]
2F3L	B1WVN5	Luminal Rfr-domain protein	<i>Cyanothece</i> sp. 51142	20	184	2	[53]
1W2I	P84142	Acylphosphatase	<i>Pyrococcus horikoshii</i>	10	91	0	[50]
4NBO	Q9HD15	Steroid receptor RNA activator protein carboxy-terminal domain	<i>Homo sapiens</i>	13	111	2	[54]
3APA	O60844	Human pancreatic secretory protein ZG16p	<i>Homo sapiens</i>	16	141	0	[55]
4IGI	E9PWQ3	Collagen VI alpha3 N5 domain	<i>Mus musculus</i>	22	203	0	[56]
2W4P	P07311	Human common-type acylphosphatase variant, A99G	<i>Homo sapiens</i>	11	99	0	[48]
4LKT	Q01469	Human Epidermal Fatty Acid Binding Protein (FABP5)	<i>Homo sapiens</i>	15	138	6	[49]
4MJJ	Q14183	C2A domain of DOC2A	<i>Homo sapiens</i>	15	138	5	To be published
4JOX	Q939T0	Cry34Ab1 protein	<i>Bacillus thuringiensis</i>	14	123	0	To be published

In five candidate proteins, we found at least one mutable surface patch, in three of them even two independent patches (Figure 1). We excluded three proteins—4NBO [54], 3APA [55], and 4IGI [56], since they did not contain any suitable surface patch.

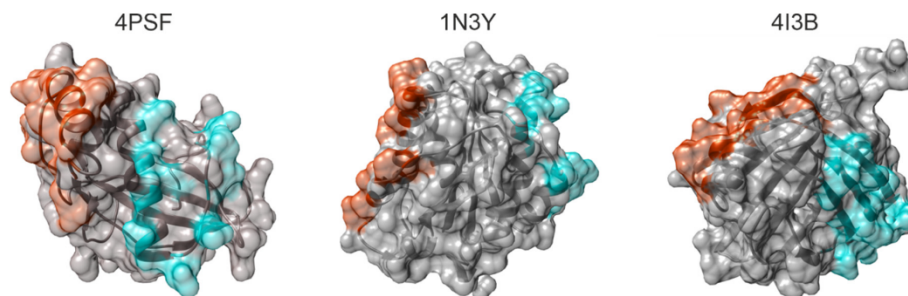


Figure 1. Surface patches of mutable residues identified in the structures of Protein Data Bank (PDB) IDs 4PSF [42], 1N3Y [52], and 4I3B [51]. The proteins are shown as ribbons in semitransparent surface representation. Residues constituting PatchN and PatchC in each protein are colored red and cyan, respectively.

3.1.3. Characterization of Soluble Wild-Type Variants of the Potential Scaffolds

We continued with five small protein scaffold candidates—4PSF [42], 1N3Y [52], 4I3B [51], 1W2I [50], and 2F3L [53]. We analyzed the expression level, solubility, purification simplicity, oligomerization state, secondary structure, and thermal stability of their wild-type (WT) variants.

(a) Protein expression and solubility. We prepared two versions of each scaffold candidate, first with N-terminal His-tag and second with C-terminal His-tag. We tested the expression and solubility of these five proteins in *E. coli* BL21(DE3) at four different temperatures, 37, 30, 25, and 16 °C. We detected high expression levels and excellent solubility of 4PSF, 1N3Y, 4I3B, and 2F3L with the N-terminal His-tag at 16 °C. The solubility varied with temperature. We were not able to express 4PSF, 1N3Y, and 4I3B proteins with the C-terminal His-tag. In the case of 1W2I, the proteins were not expressed in *E. coli* BL21(DE3) strain, only in *E. coli* C41(DE3) and C43(DE3) cells which are usually used for the expression of toxic proteins. Therefore, we decided to exclude the 1W2I protein.

(b) Protein purification. We continued with the purification of the remaining four proteins—4PSF, 1N3Y, 4I3B, and 2F3L, all of them as variants with the N-terminal His-tag. We performed a two-step purification procedure, first on Ni-NTA agarose beads and second by size exclusion chromatography. We isolated the 4PSF, 1N3Y, and 4I3B proteins at very high purity, but the 2F3L showed a minor band of either a homodimer or a non-relevant protein contaminant on SDS-PAGE (Figure 2A). The purified proteins retained solubility without aggregation and degradation at 4 °C in 20 mM sodium phosphate buffer at pH 7.5 and 100 mM NaCl in the range of days and weeks.

(c) Oligomerization. We used dynamic light scattering (DLS) to measure the oligomerization of purified proteins. The 4PSF, 1N3Y, and 4I3B proteins were monodisperse and monomeric based on the intensity-based particle size distribution with Z-averages of 4.6, 5.4, and 4.4 nm in diameter, respectively (Figure 2B). Although the 2F3L protein was monodisperse, the Z-average of 7.8 nm in diameter was higher than expected. It correlates with the impurities appearing on SDS-PAGE (Figure 2A). Since other than monomeric states of a protein would complicate biophysical measurements and binding assays, we decided to exclude 2F3L.

(d) Secondary structure. We checked the folding of the remaining three protein scaffold candidates, 4PSF, 1N3Y, and 4I3B, by circular dichroism (CD) measurements in the ultraviolet (UV) range (Figure 3). Based on the standard analysis of CD spectra using the CDNN software [59], the spectra of all three proteins showed features of regular secondary structures. In case of 4PSF and 1N3Y the combination of alpha-helices and beta-sheets, in case of 4I3B mainly of beta-sheets, agreeing qualitatively with the features of the respective crystal structures.

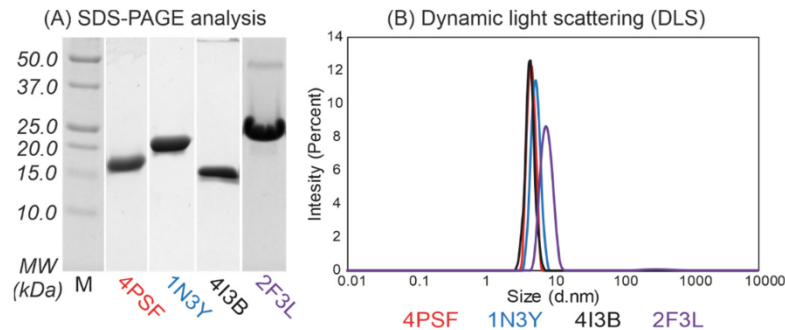


Figure 2. (A) Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE, 15%) analysis and (B) dynamic light scattering (DLS) measurement of the purified protein scaffold candidates with N-terminal His-tags after size exclusion chromatography. The 4PSF, 1N3Y, and 4I3B proteins were isolated to high-purity as monomeric proteins. On the contrary, 2F3L showed minor contaminations around 45 kDa on SDS-PAGE gel and DLS measurement confirmed the presence of bigger particles.

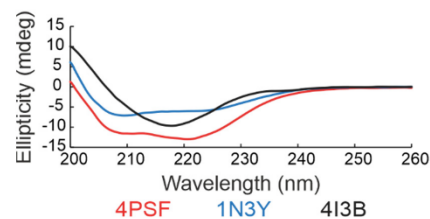


Figure 3. Far-ultraviolet (UV) circular dichroism (CD) spectra of three scaffold candidates. The results indicate that all three proteins acquire the expected secondary structure features. The 4PSF (red) and 1N3Y (blue) showed a combination of alpha-helices and beta-sheets, whereas the 4I3B (black) formed mainly beta-sheets.

(e) **Thermal stability.** We determined the thermal stability of proteins by measuring their melting temperature (T_m) using a circular dichroism technique. The melting temperatures of the WT variants of 4PSF, 1N3Y, and 4I3B were 75 °C, 57 °C, and 47 °C, respectively.

3.1.4. Testing the Stability of the Potential Scaffolds

Before starting a directed molecular evolution campaign, we experimentally verified the mutability of the predicted surface patches, for which we mutated all selected residues in each surface patch to alanines. The alanine-scanning mutagenesis is a method used to identify the structural role of protein residues [60–62]. The idea was that the protein scaffold candidate should sustain the expression, solubility, and folding upon forming a large relatively hydrophobic surface patch of alanines.

Each of three small protein scaffold candidates, 4PSF, 1N3Y, and 4I3B, has two mutable surface patches, either closer to N-terminus (PatchN) or C-terminus (PatchC) (Figure 1). Therefore, we constructed six allAla mutants. We performed the same characterization of the allAla mutants as for the WT proteins. The number of mutated residues, protein expression, solubility, and melting temperatures are summarized in Table 2. Based on these results, we decided not to continue with the 4I3B candidate because its allAla mutants did not express well or were only poorly soluble. The 1N3Y-allAla mutants were both expressed, but only the PatchN was soluble. The best results were achieved for 4PSF because both its allAla-PatchN and allAla-PatchC variants were expressed and soluble.

Table 2. Properties of three most promising wild-type (WT) protein scaffold candidates and their allAla variants. The melting temperatures (T_m) were measured by circular dichroism.

Protein Scaffold Candidate Variant	Number of Mutated Residues	Expression	Solubility	T_m
4PSF-WT	0	Yes	Very good	75 °C
4PSF-allAla-PatchN	10	Yes	Very good	60 °C
4PSF-allAla-PatchC	10	Yes	Very good	67 °C
1N3Y-WT	0	Yes	Very good	57 °C
1N3Y-allAla-PatchN	10	Yes	Good	51 °C
1N3Y-allAla-PatchC	11	Yes	Not soluble	N/A
4I3B-WT	0	Yes	Good	47 °C
4I3B-allAla-PatchN	12	No	No expression	N/A
4I3B-allAla-PatchC	11	Yes	Poor	N/A

Next, we performed control measurements to check for non-specific binding of the 1N3Y and 4PSF scaffold candidates and their allAla mutants. We measured their binding to the capillaries used for microscale thermophoresis (MST) and to sensor chips for surface plasmon resonance (SPR) (Figure S2). We determined that the 1N3Y-WT and 1N3Y-allAla-PatchN were applicable for MST but not for SPR because they non-specifically bind to a clean surface of the SPR sensor chip. We observed no such effects with any of the 4PSF variants.

Considering all the findings, we chose the 4PSF protein as the best scaffold candidate. The WT protein as well as the allAla mutants were expressed at high yield, soluble and easy to purify, were monodispersed and folded, and possessed appropriate characteristics for MST and SPR affinity measurements. We selected the 4PSF C-terminal surface patch (PatchC) for further work because the 4PSF-PatchC-allAla mutant had the highest melting temperature among the allAla mutants.

3.1.5. Testing the Suitability of the Scaffold Candidates for In Vitro Evolution

First, we constructed a degenerate DNA library for ribosome display. The selected PatchC of the 4PSF scaffold candidate had 10 calculated mutable residues (Figure 4). The random mutagenesis of these 10 positions gave rise to the maximal diversity of more than 1012 variants, which agreed with the potential diversity suitable for ribosome display [57,58]. The DNA library was synthesized by GENEWIZ company by the degenerate NNK codons technology. The same company estimated library correctness as being 78%, which was sufficient for further work.

Next, we performed pilot ribosome display selection and crystallization of variants. We took advantage of directed evolution to train the 4PSF scaffold candidate to bind the IL-10 protein as the target molecule. Pilot experiments were undertaken to establish the ribosome display protocol for the initial preselection process (see section Section 3.2.2 below). After the fourth selection round of ribosome display, we selected eight random colonies for two-step purification and crystallization trials. We used commercial screens for preliminary crystallization trials and then manually optimized the conditions to get better diffracting crystals. We started the crystallization with seven variants (G14, G21, H25, H33, J61, J70, and J93). We observed a crystallization process of six variants, and only the G21 did not result in any crystal form in our tested screens. Four variants created crystals with low diffraction quality (ranging 6 to 8 Å). Crystallization of the J61 variant resulted in crystals diffracting up to a high resolution (1.2 Å). We have solved the structure with the molecular replacement method using the original 4PSF structure [42]. The structure, deposited under the PDB entry ID 7AVC, was similar to the original 4PSF structure with root mean square deviation of 0.93 Å calculated on 129 C α atoms (Figure S3). The fold remains stable even after nine mutations at the selected positions. Thus, we established the suitability of the 4PSF scaffold candidate for in vitro evolution by ribosome display. Because the affinity

of the J61 variant to IL-10 was only in the micromolar range we continued in the effort to develop higher affinity binders.

Variant	Position																									
	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134
WT	R	E	L	V	I	T	I	A	R	E	G	L	E	D	K	Y	N	L	Q	L	N	P	E	W	R	M
A2	R	G	L	V	I	R	I	A	Q	R	G	I	E	F	K	Y	L	L	A	L	N	P	R	W	I	M
A3	R	P	L	V	I	R	I	A	V	G	G	L	E	R	K	Y	G	L	S	L	P	P	L	W	R	M
B4	R	R	L	V	I	R	I	A	R	T	G	L	E	L	K	Y	S	L	N	L	W	P	P	W	S	M
C4	R	V	L	V	I	L	I	A	L	L	G	L	E	V	K	Y	R	L	A	L	Q	P	V	W	Y	M
C11	R	L	L	V	I	L	I	A	I	V	G	L	E	W	K	Y	P	L	P	L	V	P	L	W	E	M
C12	R	G	L	V	I	E	I	A	P	T	G	L	E	W	K	Y	F	L	L	L	E	P	S	W	C	M
E3	R	G	L	V	I	G	I	A	H	R	G	L	E	S	K	Y	Y	L	R	L	G	P	R	W	W	M
F5	R	R	L	V	I	T	I	A	L	R	G	L	E	L	K	Y	P	L	C	L	R	P	A	W	H	M
G3	R	R	L	V	I	A	I	A	P	N	G	L	E	R	K	Y	T	L	H	L	T	P	T	W	S	M
G6	R	G	L	V	I	F	I	A	V	G	G	L	E	S	K	Y	L	L	D	L	E	P	L	W	H	M

Figure 4. The amino acid sequences of the wild-type (WT) 4PSF protein scaffold and its selected variants with the highest affinity to interleukin-10 and the lowest affinity to interleukin-29, bovine serum albumin (BSA), and skimmed milk. The yellow color represents amino acids of 4PSF-WT protein that were mutated (in green) in scaffold variants. The residue numbering corresponds to the PDB entry 4PSF.

3.1.6. Construction of the Final ProBi Scaffold

The final 4PSF scaffold construct used for directed evolution experiments, termed ProBi-WT hereafter, consists of three sequence segments (Figure 5): (i) the N-terminal Strep-tag to simplify purification; (ii) the 4PSF scaffold segment; (iii) the C-terminal c-Myc-tag. It was added to detect variants expressed properly to the c-Myc tag and exclude fragments emerging because of premature stop codons or frameshifts in the randomized DNA library. The expression level, solubility, and biophysical features of this new ProBi scaffold construct were similar to the original scaffold candidate.

MAW*SHPQFEK SMAQGPGQPG FCIKTN***SSEG** *KVFINICHSP SIPPPADVTE EELLQMLEED*
AGFRIPMSL *GEPHAELDAK GQGCTAYDVA VNSDFYRRMQ NSDFLRELVI* **TIAREGLEDK**
YNLQLNPEWR *MKNRPFMGS IGSEQKLI***SE EDL**

Figure 5. The amino acid sequence of the wild type ProBi-WT protein scaffold. The N-terminal Strep-tag and C-terminal c-Myc-tag are in bold, the 4PSF scaffold segment is in italic. The ten mutable amino acid residues forming the PatchC are highlighted in red.

3.2. Application of the ProBi Scaffold on the Interleukin-10 (IL-10) Target as a Model System

3.2.1. Selection of the Target Protein

Cytokines from the family of interleukin-10 are medically important signaling proteins of native immunity [31–33]. Immunosuppressive effects of the prototypical member of the family, IL-10, which may hinder immunotherapeutic strategies of cancer therapy [32], are not completely understood. In addition, monodispersity, CD spectra, and signaling on the RAW264.7 murine cell line of recombinant IL-10 produced in the laboratory encouraged use of this cytokine as the biologically relevant and molecularly well-defined target for our study.

3.2.2. Selection of Novel ProBi-Based Binders by Ribosome Display

(a) Initial preselection process. We performed four selection rounds of ribosome display, with the following preselection conditions: incubation for an hour with a mixture of BSA, IFN γ , and ProBi-WT scaffold; blocking solution 3% BSA for all rounds. After the fourth round, we selected 190 random colonies (together with ProBi-WT) for a small-scale expression in deep-96-well plates and for detection of binding to IL-10 using the ELISA assay. We took advantage of the C-terminal c-Myc-tag to reveal those clones that have the correct open reading frame and bind the IL-10. From these, we selected 47 clones with the highest signal for larger-scale expression in deep-24-well plates. These 47 clones were tested for binding to IL-29, which shares the IL-10R2 receptor with IL-10 [63]. We discovered that most of the clones showed binding to IL-29 as well as to IL-10. Most variants also displayed non-specific binding to BSA and skimmed milk.

(b) Optimized preselection process. To decrease the non-specific binding of ProBi variants to BSA and milk proteins, we optimized the preselection conditions. The initial preselection conditions were extended in three ways. We (1) included IL-29 to the preselection mix; (2) varied the blocking solution between BSA (used in the first and third round) and skimmed milk (used in the second and fourth round); and (3) performed the preselection reaction in three consecutive wells. Otherwise, we completed the ribosome display the same way as in the point a) above. The optimization of the preselection monitored on 190 variants led to a decrease of non-specific binding to the BSA and milk to 50% and 60%, respectively, of the binding to IL-10. For further characterization, we selected 10 variants (Figure 4) with high affinities to IL-10 and low affinities to IL-29, BSA, and milk. We created the phylogenetic tree that is described in Supporting Information (Figure S4).

3.2.3. Characterization of Novel ProBi-Based Binders

(a) Affinity estimation by ELISA assay. We expressed and purified ten selected variants by one step purification on affinity StrepTactinXT resin and performed ELISA measurements to estimate their affinities to IL-10 and BSA (data not shown). We discarded six variants with the highest binding background to BSA and continued with four variants for more detailed characterization.

(b) Affinity measurements by MST. We expressed and purified ProBi-WT and its four variants labeled F5, G3, A2, and G6 by two-step purification on affinity StrepTactinXT resin and size exclusion chromatography. We successfully measured the IL-10 binding affinities of two variants, G3 and F5 (Figure 6). Their dissociation constants (K_d) are shown in Table 3. The F5 variant had the highest affinities, 6 nM in Tris (pH 8.0) buffer, and 15 nM in Hepes (pH 7.5) buffer, respectively. We detected binding of the A2 variant to IL-10 but we were not able to reach the bound (saturation) state of the binding curve. Therefore, we only estimated the affinities to be higher than 1 μ M. The G3 variant assembled into a dimer, and the A2 and F5 variants behaved the same way as the ProBi-WT protein. The G6 variant showed the highest tendency to form oligomers (data not shown) and, therefore, it was discarded.

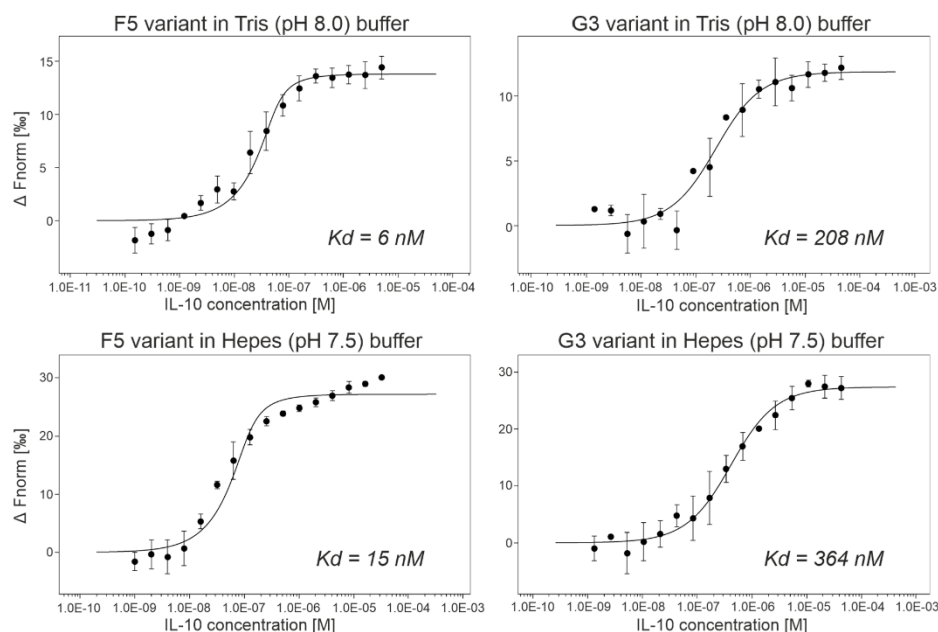


Figure 6. The microscale thermophoresis (MST) curves of two best ProBi IL-10 binders, F5 and G3. The results showed that the change of buffer from Tris to Hepes and pH from 8.0 to 7.5 did not affect the affinities.

Table 3. Variants of the ProBi scaffold sorted by their affinities to IL-10. The table shows affinities, response amplitude during microscale thermophoresis (MST) measurements, number of replicates of the MST measurements, and melting temperatures (T_m) measured by the nanoDSF method. Not measurable indicates that the affinity is so low that it falls beyond the dynamic range of the MST technique.

ProBi Variant (Buffer)	Affinity	Response Amplitude	Replicates	T_m
WT (Tris)	Not measurable	Not measurable	3	66 °C
F5 (Tris)	6 nM	14	4	51 °C
F5 (Hepes)	15 nM	27	3	N/A
G3 (Tris)	208 nM	12	3	58 °C
G3 (Hepes)	364 nM	27	3	N/A
A2 (Tris)	>1 μ M	N/A	2	48 °C

To evaluate the affinities of ProBi variants to IL-10, we utilized the microscale thermophoresis technique that utilizes low material consumption [64]. We used a commercial kit to fluorescently label the IL-10 with C-terminal His-tag as a target molecule and titrated it by evolved ProBi variants with the N-terminal Strep-tag. The measured affinities are shown in Table 3. We did not detect binding of IL-10 to the parental wild-type ProBi scaffold in our binding buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.1% Pluronic F-127). We cannot measure the affinities to BSA and milk in our MST setup because they do not possess His-tag for labeling.

(c) Characterization of ProBi variants with the highest affinity. The chromatograms from size exclusion chromatography and SDS-PAGE analysis of purified proteins are shown in Figure S5. We measured the melting temperature (T_m) of the variants (Table 3) finding that T_m was lower than for the WT protein, but still in a range acceptable for practical

purposes. We confirmed the folded structure of the selected ProBi variants with a high content of alpha helices according to their circular dichroism spectra (Figure S6).

(d) Inhibitory assay. We tested potential of the G3 and F5 variants to inhibit the IL-10 signaling pathway by a competitive binding assay on the RAW264.7 cell line, which expresses both IL10R1 and IL10R2 receptors on the cell surface. We used ProBi-WT protein as a negative control. Using these experimental conditions, we observed no inhibition of the IL-10 signaling pathway by either G3 or F5 (Figure S7). We hypothesize that these ProBi variants bind to the surface of IL-10 in such a way that they do not prevent IL-10 binding to the receptors IL10R1 and/or IL10R2 and do not block the signalization.

In order to identify truly inhibitory binders with a potential medical use, we need to include more ProBi variants, correlate their activity with the effect of neutralizing anti-IL-10 antibodies, test the inhibition on more cell types (e.g., THP-1 or U937), and possibly monitor signalization by outcomes other than STAT3 phosphorylation. The ProBi IL-10 binders developed as a proof of principle will have to be engineered further in order to inhibit the signaling pathway of IL-10 cytokine.

3.2.4. Further Development of Scaffold Selection and Improvement of Probi-Based Binders

We are aware that the procedure of scaffold selection can be optimized, and here we discuss possible future modifications of the present protocol.

(a) The first step of our procedure, the visual screening of the PDB, can be performed more systematically. A possible way would be to automate selection of monomeric small proteins and the following identification of mutable patches on their surfaces by in silico screening by employing the FoldX [36] or Rosetta [65] programs. We can also include proteins that were produced not only in *E. coli* expression systems because the expression level and solubility of the tested scaffolds can be tested in a high-throughput format.

(b) The proposed workflow will be further tested by development of our second candidate scaffold, protein alpha-X beta2 integrin I domain (UniProt ID P20702, structure of PDB ID 1N3Y [52]) in the near future.

(c) The selectivity of the best ProBi binders to other IL-10 family cytokines was tested only by the ELISA pre-screening against IL-29 (Section 3.2.2). Further development of the binders would undoubtedly require thorough checking of their cross-reactivity to more cytokines and other proteins.

(d) A substantial challenge for possible medical use of the binders is their immunogenicity [24]. We do not expect the ProBi-based binders to be immunogenic as the scaffold is derived from the human PIH1 domain-containing protein 1 (UniProt ID Q9NWS0, structure of PDB ID 4PSF [42]). However, potentially immunogenic scaffolds could be repurposed for molecular imaging [16].

(e) The directed evolution methodologies [66,67] offer avenues to development of proteins with new properties. Over the years, several display techniques [68,69] have evolved and they offer alternative and complementary ways to the selection of optimal protein molecules for the task. In this work, we make use only one of the display techniques, ribosome display. Because the yeast display offers certain advantages compared to ribosome display, we designed the final ProBi scaffold construct to be directly used by both methods. Therefore, we included the C-terminal c-Myc-tag that is widely used for high-throughput detection in Fluorescence-activated cell sorting (FACS).

(f) The protein scaffolds as well as antibodies typically bind just one interaction partner. Two independently mutable patches on one protein scaffold molecule could open a way to train binders to interact simultaneously with two partners. Therefore, we aimed at scaffolds with two mutable patches. Such a feature would mimic a function of natural proteins, such as cytokines and/or other proteins signaling through simultaneous binding to two receptors i.e., via formation of a ternary complex. The protein scaffolds can also work as “synthekines” to engineer non-natural receptor heterodimers that could activate new unexplored cellular responses [70].

4. Conclusions

Protein scaffolds represent a great engineering tool that could complement functions of more commonly used antibodies in high-affinity binding of biomolecular targets. We think that there is no “one-scaffold-fits-all” and that the development of new scaffolds tailored to specific functions is crucial for real-life applications, e.g., regulation of signaling pathways *in vivo*, biologics, and molecular imaging.

Visual inspection of several hundred structures from the PDB, *in silico* investigation of about 100 selected candidates, and finally experimental examination of 12 of them, led to selection of two potentially useful scaffold proteins. The first candidate was based on the human PIH1 domain-containing protein 1 (UniProt ID Q9NWS0, structure of PDB ID 4PSF [42]) and the second one on the alpha-X beta2 integrin I domain (UniProt ID P20702, structure of PDB ID 1N3Y [52]). The latter of the two candidates is going to be developed further in the near future.

We preferred the 4PSF protein structure because it contained two surface patches amenable for independent mutagenesis. Thus, we modified the 4PSF protein construct into the protein scaffold construct called ProBi that was ready to be used in both ribosome and yeast display technology. For the purpose of this work, we focused on one of the patches and utilized the ribosome display to evolve the ProBi scaffold into the binders of interleukin-10 with nanomolar affinity. In the future, we plan to include more scaffold candidates and concentrate on their other features, such as signaling inhibition, immunogenicity, and selectivity.

In this work, we present a proof of concept methodology to identify protein structures that could be converted into a new protein scaffold constructs. We experimentally proved that at least one of them could be adapted into binders of a medically important target with nanomolar affinity by methods of directed evolution.

Supplementary Materials: The following are available online at <https://www.mdpi.com/1999-4915/13/2/190/s1>, Table S1: List of primers, Table S2: Data processing statistics of PDB entry 7AVC. Figure S1: DNA sequences of internal arrangement within (A) the pRDVsm and (B) pETsm vectors. Figure S2: Testing of non-specific binding of scaffold candidates. Figure S3: Structural alignment. Figure S4: The phylogenetic tree of ten ProBi scaffold variants. Figure S5: SDS-PAGE (15%) analysis of purified ProBi variants. Figure S6: Circular dichroism spectra of (A) ProBi scaffold. Figure S7: Inhibition of the IL-10 signaling pathway estimated by a competitive binding assay.

Author Contributions: Conceptualization, B.S. and P.M.; methodology, B.S., P.M. and J.Z.; computational analysis, L.B., J.Č. and B.S.; molecular biology experiments, P.N.P., M.H., L.K. and P.M.; biophysical experiments P.N.P., M.H., T.C., G.F., P.K., Š.H. and J.P.; data analysis, M.H., P.N.P., Š.H., L.K., P.K., J.P. and J.Z.; writing—original draft preparation, B.S., P.M. and L.B.; writing—review and editing, B.S., P.M., P.N.P. and M.H.; visualization, P.M., M.H. and P.N.P.; supervision, B.S.; project administration, B.S. and P.M.; funding acquisition, B.S.; P.N.P. and M.H. contributed equally to the manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Czech Science Foundation, grant 19-17398S, Czech Academy of Sciences, grant RVO 86652036, Ministry of Education, Youth, and Sports of the Czech Republic, grant LM2015043, LM2018127.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The structural data presented in this study are openly available in the Protein Data Bank, identification number 7AVC. All other data presented in this study are available on request from the corresponding authors.

Acknowledgments: We thank Marketa Janovska for her invaluable technical assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bryson, J.W.; Desjarlais, J.R.; Handel, T.M.; DeGrado, W.F. From coiled coils to small globular proteins: Design of a native-like three-helix bundle. *Protein Sci.* **1998**, *7*, 1404–1414. [[CrossRef](#)] [[PubMed](#)]
2. Khoury, G.A.; Smadbeck, J.; Kieslich, C.A.; Floudas, C.A. Protein folding and de novo protein design for biotechnological applications. *Trends Biotechnol.* **2014**, *32*, 99–109. [[CrossRef](#)] [[PubMed](#)]
3. Burnside, D.; Schoenrock, A.; Moteshareie, H.; Hooshyar, M.; Basra, P.; Hajikarimlou, M.; Dick, K.; Barnes, B.; Kazmirchuk, T.; Jessulat, M.; et al. In Silico Engineering of Synthetic Binding Proteins from Random Amino Acid Sequences. *iScience* **2019**, *11*, 375–387. [[CrossRef](#)] [[PubMed](#)]
4. Kortemme, T.; Baker, D. Computational design of protein-protein interactions. *Curr. Opin. Chem. Biol.* **2004**, *8*, 91–97. [[CrossRef](#)] [[PubMed](#)]
5. Mikulecky, P.; Cerny, J.; Biedermannova, L.; Petrokova, H.; Kuchar, M.; Vondrasek, J.; Maly, P.; Sebo, P.; Schneider, B. Increasing affinity of interferon-gamma receptor 1 to interferon-gamma by computer-aided design. *BioMed Res. Int.* **2013**, *2013*, 752514. [[CrossRef](#)]
6. Smith, G.P.; Petrenko, V.A. Phage Display. *Chem. Rev.* **1997**, *97*, 391–410. [[CrossRef](#)]
7. Hosse, R.J.; Rothe, A.; Power, B.E. A new generation of protein display scaffolds for molecular recognition. *Protein Sci.* **2006**, *15*, 14–27. [[CrossRef](#)]
8. Brender, J.R.; Shultis, D.; Khattak, N.A.; Zhang, Y. An Evolution-Based Approach to De Novo Protein Design. *Methods Mol. Biol.* **2017**, *1529*, 243–264.
9. Goldenzweig, A.; Goldsmith, M.; Hill, S.E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; et al. Automated Structure- and Sequence-Based Design of Proteins for High Bacterial Expression and Stability. *Mol. Cell* **2016**, *63*, 337–346. [[CrossRef](#)]
10. Musil, M.; Stourac, J.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Zendulka, J.; Martinek, T.; Bednar, D.; Damborsky, J. FireProt: Web server for automated design of thermostable proteins. *Nucleic Acids Res.* **2017**, *45*, W393–W399. [[CrossRef](#)]
11. Zahradnik, J.; Kolarova, L.; Peleg, Y.; Kolenko, P.; Svidenska, S.; Chamavets, T.; Unger, T.; Sussman, J.L.; Schneider, B. Flexible regions govern promiscuous binding of IL-24 to receptors IL-20R1 and IL-22R1. *FEBS J.* **2019**, *286*, 3858–3873. [[CrossRef](#)] [[PubMed](#)]
12. Banta, S.; Dooley, K.; Shur, O. Replacing antibodies: Engineering new binding proteins. *Annu. Rev. Biomed. Eng.* **2013**, *15*, 93–113. [[CrossRef](#)]
13. Yu, X.; Yang, Y.P.; Dikici, E.; Deo, S.K.; Daunert, S. Beyond Antibodies as Binding Partners: The Role of Antibody Mimetics in Bioanalysis. *Annu. Rev. Anal. Chem.* **2017**, *10*, 293–320. [[CrossRef](#)] [[PubMed](#)]
14. Skerra, A. Engineered protein scaffolds for molecular recognition. *J. Mol. Recognit.* **2000**, *13*, 167–187. [[CrossRef](#)]
15. Hey, T.; Fiedler, E.; Rudolph, R.; Fiedler, M. Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends Biotechnol.* **2005**, *23*, 514–522. [[CrossRef](#)]
16. Simeon, R.; Chen, Z. In vitro-engineered non-antibody protein therapeutics. *Protein Cell* **2018**, *9*, 3–14. [[CrossRef](#)]
17. Nord, K.; Gunneriusson, E.; Ringdahl, J.; Stahl, S.; Uhlen, M.; Nygren, P.A. Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat. Biotechnol.* **1997**, *15*, 772–777. [[CrossRef](#)]
18. Binz, H.K.; Amstutz, P.; Pluckthun, A. Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol.* **2005**, *23*, 1257–1268. [[CrossRef](#)]
19. Skerra, A. Alternative non-antibody scaffolds for molecular recognition. *Curr. Opin. Biotechnol.* **2007**, *18*, 295–304. [[CrossRef](#)]
20. Barinka, C.; Ptacek, J.; Richter, A.; Novakova, Z.; Morath, V.; Skerra, A. Selection and characterization of Anticalins targeting human prostate-specific membrane antigen (PSMA). *Protein Eng. Des. Select.* **2016**, *29*, 105–115. [[CrossRef](#)]
21. Bedford, R.; Tiede, C.; Hughes, R.; Curd, A.; McPherson, M.J.; Peckham, M.; Tomlinson, D.C. Alternative reagents to antibodies in imaging applications. *Biophys. Rev.* **2017**, *9*, 299–308. [[CrossRef](#)] [[PubMed](#)]
22. Rothe, C.; Skerra, A. Anticalin((R)) Proteins as Therapeutic Agents in Human Diseases. *BioDrugs* **2018**, *32*, 233–243. [[CrossRef](#)] [[PubMed](#)]
23. Kosztyu, P.; Kuchar, M.; Cerny, J.; Barkocziava, L.; Maly, M.; Petrokova, H.; Czernekova, L.; Liskova, V.; Raskova Kafkova, L.; Knotigova, P.; et al. Proteins mimicking epitope of HIV-1 virus neutralizing antibody induce virus-neutralizing sera in mice. *EBioMedicine* **2019**, *47*, 247–256. [[CrossRef](#)] [[PubMed](#)]
24. Vazquez-Lombardi, R.; Phan, T.G.; Zimmermann, C.; Lowe, D.; Jermutus, L.; Christ, D. Challenges and opportunities for non-antibody scaffold drugs. *Drug Discov. Today* **2015**, *20*, 1271–1283. [[CrossRef](#)] [[PubMed](#)]
25. Azhar, A.; Ahmad, E.; Zia, Q.; Rauf, M.A.; Owais, M.; Ashraf, G.M. Recent advances in the development of novel protein scaffolds based therapeutics. *Int. J. Biol. Macromol.* **2017**, *102*, 630–641. [[CrossRef](#)]
26. Ahmad, J.N.; Li, J.; Biedermannova, L.; Kuchar, M.; Sipova, H.; Semeradtova, A.; Cerny, J.; Petrokova, H.; Mikulecky, P.; Polinek, J.; et al. Novel high-affinity binders of human interferon gamma derived from albumin-binding domain of protein G. *Proteins* **2012**, *80*, 774–789. [[CrossRef](#)]
27. Kuchar, M.; Vankova, L.; Petrokova, H.; Cerny, J.; Osicka, R.; Pelak, O.; Sipova, H.; Schneider, B.; Homola, J.; Sebo, P.; et al. Human interleukin-23 receptor antagonists derived from an albumin-binding domain scaffold inhibit IL-23-dependent ex vivo expansion of IL-17-producing T-cells. *Proteins* **2014**, *82*, 975–989. [[CrossRef](#)]

28. Kruziki, M.A.; Bhatnagar, S.; Woldring, D.R.; Duong, V.T.; Hackel, B.J. A 45-Amino-Acid Scaffold Mined from the PDB for High-Affinity Ligand Engineering. *Chem. Biol.* **2015**, *22*, 946–956. [[CrossRef](#)]
29. Alsultan, A.M.; Chin, D.Y.; Howard, C.B.; de Bakker, C.J.; Jones, M.L.; Mahler, S.M. Beyond Antibodies: Development of a Novel Protein Scaffold Based on Human Chaperonin 10. *Sci. Rep.* **2016**, *5*, 37348. [[CrossRef](#)]
30. Berman, H.M.; Battistuzzi, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58 Pt 6*, 899–907. [[CrossRef](#)]
31. El Kasmi, K.C.; Smith, A.M.; Williams, L.; Neale, G.; Panopoulos, A.D.; Watowich, S.S.; Hacker, H.; Foxwell, B.M.; Murray, P.J. Cutting edge: A transcriptional repressor and corepressor induced by the STAT3-regulated anti-inflammatory signaling pathway. *J. Immunol.* **2007**, *179*, 7215–7219. [[CrossRef](#)] [[PubMed](#)]
32. Asadullah, K.; Sterry, W.; Volk, H.D. Interleukin-10 therapy—Review of a new approach. *Pharmacol. Rev.* **2003**, *55*, 241–269. [[CrossRef](#)] [[PubMed](#)]
33. Liu, Y.; de Waal Malefyt, R.; Briere, F.; Parham, C.; Bridon, J.M.; Banchereau, J.; Moore, K.W.; Xu, J. The EBV IL-10 homologue is a selective agonist with impaired binding to the IL-10 receptor. *J. Immunol.* **1997**, *158*, 604–613. [[PubMed](#)]
34. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)] [[PubMed](#)]
35. Okonechnikov, K.; Golosova, O.; Fursov, M.; Ugene Team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [[CrossRef](#)] [[PubMed](#)]
36. Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX web server: An online force field. *Nucleic Acids Res.* **2005**, *33*, W382–W388. [[CrossRef](#)]
37. Mikulecky, P.; Zahradnik, J.; Kolenko, P.; Cerny, J.; Charnavets, T.; Kolarova, L.; Necasova, I.; Pham, P.N.; Schneider, B. Crystal structure of human interferon-gamma receptor 2 reveals the structural basis for receptor specificity. *Acta Crystallogr. D Struct. Biol.* **2016**, *72 Pt 9*, 1017–1025. [[CrossRef](#)]
38. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 2*, 125–132. [[CrossRef](#)]
39. Evans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **2013**, *69 Pt 7*, 1204–1214. [[CrossRef](#)]
40. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67 Pt 4*, 235–242. [[CrossRef](#)]
41. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40 Pt 4*, 658–674. [[CrossRef](#)]
42. Horejsi, Z.; Stach, L.; Flower, T.G.; Joshi, D.; Flynn, H.; Skehel, J.M.; O'Reilly, N.J.; Ogdrowicz, R.W.; Smerdon, S.J.; Boulton, S.J. Phosphorylation-dependent PIH1D1 interactions define substrate specificity of the R2TP cochaperone complex. *Cell Rep.* **2014**, *7*, 19–26. [[CrossRef](#)] [[PubMed](#)]
43. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67 Pt 4*, 355–367. [[CrossRef](#)]
44. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66 Pt 4*, 486–501. [[CrossRef](#)]
45. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293–315. [[CrossRef](#)] [[PubMed](#)]
46. Dreier, B.; Pluckthun, A. Ribosome display: A technology for selecting and evolving proteins from large libraries. *Methods Mol. Biol.* **2011**, *687*, 283–306. [[PubMed](#)]
47. Zahnd, C.; Amstutz, P.; Pluckthun, A. Ribosome display: Selecting and evolving proteins in vitro that specifically bind to a target. *Nat. Methods* **2007**, *4*, 269–279. [[CrossRef](#)]
48. Lam, S.Y.; Yeung, R.C.; Yu, T.H.; Sze, K.H.; Wong, K.B. A rigidifying salt-bridge favors the activity of thermophilic enzyme at high temperatures at the expense of low-temperature activity. *PLoS Biol.* **2011**, *9*, e1001027. [[CrossRef](#)]
49. Armstrong, E.H.; Goswami, D.; Griffin, P.R.; Noy, N.; Ortlund, E.A. Structural basis for ligand regulation of the fatty acid-binding protein 5, peroxisome proliferator-activated receptor beta/delta (FABP5-PPARbeta/delta) signaling pathway. *J. Biol. Chem.* **2014**, *289*, 14941–14954. [[CrossRef](#)]
50. Cheung, Y.Y.; Lam, S.Y.; Chu, W.K.; Allen, M.D.; Bycroft, M.; Wong, K.B. Crystal structure of a hyperthermophilic archaeal acylphosphatase from *Pyrococcus horikoshii*—Structural insights into enzymatic catalysis, thermostability, and dimerization. *Biochemistry* **2005**, *44*, 4601–4611. [[CrossRef](#)]
51. Kumagai, A.; Ando, R.; Miyatake, H.; Greimel, P.; Kobayashi, T.; Hirabayashi, Y.; Shimogori, T.; Miyawaki, A. A bilirubin-inducible fluorescent protein from eel muscle. *Cell* **2013**, *153*, 1602–1611. [[CrossRef](#)] [[PubMed](#)]
52. Vorup-Jensen, T.; Ostermeier, C.; Shimaoka, M.; Hommel, U.; Springer, T.A. Structure and allosteric regulation of the alpha X beta 2 integrin I domain. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1873–1878. [[CrossRef](#)] [[PubMed](#)]

53. Buchko, G.W.; Ni, S.; Robinson, H.; Welsh, E.A.; Pakrasi, H.B.; Kennedy, M.A. Characterization of two potentially universal turn motifs that shape the repeated five-residues fold—Crystal structure of a luminal pentapeptide repeat protein from Cyanothoece 51142. *Protein Sci.* **2006**, *15*, 2579–2595. [[CrossRef](#)] [[PubMed](#)]
54. McKay, D.B.; Xi, L.; Barthel, K.K.B.; Cech, T.R. Structure and function of steroid receptor RNA activator protein, the proposed partner of SRA noncoding RNA. *J. Mol. Biol.* **2014**, *426*, 1766–1785. [[CrossRef](#)] [[PubMed](#)]
55. Kanagawa, M.; Satoh, T.; Ikeda, A.; Nakano, Y.; Yagi, H.; Kato, K.; Kojima-Aikawa, K.; Yamaguchi, Y. Crystal structures of human secretory proteins ZG16p and ZG16b reveal a Jacalin-related beta-prism fold. *Biochem. Biophys. Res. Commun.* **2011**, *404*, 201–205. [[CrossRef](#)]
56. Becker, A.K.; Mikolajek, H.; Paulsson, M.; Wagener, R.; Werner, J.M. A structure of a collagen VI VWA domain displays N and C termini at opposite sides of the protein. *Structure* **2014**, *22*, 199–208. [[CrossRef](#)]
57. Groves, M.A.; Osbourn, J.K. Applications of ribosome display to antibody drug discovery. *Expert Opin. Biol. Ther.* **2005**, *5*, 125–135. [[CrossRef](#)]
58. Pluckthun, A. Ribosome display: A perspective. *Methods Mol. Biol.* **2012**, *805*, 3–28.
59. Bohm, G.; Muhr, R.; Jaenicke, R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **1992**, *5*, 191–195. [[CrossRef](#)] [[PubMed](#)]
60. Lefevre, F.; Remy, M.H.; Masson, J.M. Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function. *Nucleic Acids Res.* **1997**, *25*, 447–448. [[CrossRef](#)] [[PubMed](#)]
61. Weiss, G.A.; Watanabe, C.K.; Zhong, A.; Goddard, A.; Sidhu, S.S. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8950–8954. [[CrossRef](#)] [[PubMed](#)]
62. Morrison, K.L.; Weiss, G.A. Combinatorial alanine-scanning. *Curr. Opin. Chem. Biol.* **2001**, *5*, 302–307. [[CrossRef](#)]
63. Miknis, Z.J.; Magracheva, E.; Li, W.; Zdanov, A.; Kotenko, S.V.; Wlodawer, A. Crystal structure of human interferon-lambda1 in complex with its high-affinity receptor interferon-lambdaR1. *J. Mol. Biol.* **2010**, *404*, 650–664. [[CrossRef](#)] [[PubMed](#)]
64. Jerabek-Willemsen, M.; Wienken, C.J.; Braun, D.; Baaske, P.; Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev. Technol.* **2011**, *9*, 342–353. [[CrossRef](#)]
65. Moretti, R.; Lyskov, S.; Das, R.; Meiler, J.; Gray, J.J. Web-accessible molecular modeling with Rosetta: The Rosetta Online Server that Includes Everyone (ROSIE). *Protein Sci.* **2018**, *27*, 259–268. [[CrossRef](#)]
66. Hibbert, E.G.; Dalby, P.A. Directed evolution strategies for improved enzymatic performance. *Microb. Cell Factories* **2005**, *4*, 29. [[CrossRef](#)]
67. Bloom, J.D.; Arnold, F.H. In the light of directed evolution: Pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. USA* **2009**, *106* (Suppl. 1), 9995–10000. [[CrossRef](#)]
68. Packer, M.S.; Liu, D.R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **2015**, *16*, 379–394. [[CrossRef](#)]
69. Xiao, H.; Bao, Z.; Zhao, H. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **2015**, *54*, 4011–4020. [[CrossRef](#)]
70. Moraga, I.; Spangler, J.B.; Mendoza, J.L.; Gakovic, M.; Wehrman, T.S.; Krutzik, P.; Garcia, K.C. Synthekines are surrogate cytokine and growth factor agonists that compel signaling through non-natural receptor dimers. *eLife* **2017**, *6*, e22882. [[CrossRef](#)]



Diffraction anisotropy and paired refinement: crystal structure of H33, a protein binder to interleukin 10

Petr Kolenko,^{a,b*} Pavel Mikulecký,^b Phuong Ngoc Pham,^b Martin Malý^{a,b} and Bohdan Schneider^b

Received 31 October 2022
Accepted 1 June 2023

^aCzech Technical University in Prague, Břehova 7, Prague 115 19, Czech Republic, and ^bInstitute of Biotechnology of the Czech Academy of Sciences, Biocev, Průmyslová 595, Vestec 25250, Czech Republic. *Correspondence e-mail: petr.kolenko@ijji.cvut.cz

Edited by A. R. Pearson, Universität Hamburg, Germany

Keywords: anisotropy; paired refinement; binder H33.

PDB reference: 8bxu

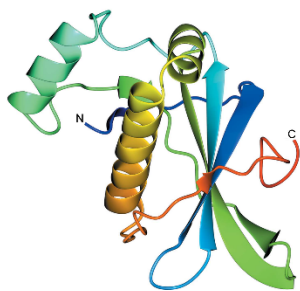
Supporting information: this article has supporting information at journals.iucr.org/j

Binder H33 is a small protein binder engineered by ribosome display to bind human interleukin 10. Crystals of binder H33 display severe diffraction anisotropy. A set of data files with correction for diffraction anisotropy based on different local signal-to-noise ratios was prepared. Paired refinement was used to find the optimal anisotropic high-resolution diffraction limit of the data: 3.13–2.47 Å. The structure of binder H33 belongs to the 2% of crystal structures with the highest solvent content in the Protein Data Bank.

1. Introduction

The diffraction quality of a crystal is usually different in various reciprocal space directions. Diffraction anisotropy can be caused by crystal growth, the crystal shape, the modulated volume of the irradiated crystal during the measurement and the arrangement of molecules inside the crystal. This phenomenon is often not a serious issue for a successful structure determination. Most of the macromolecular refinement programs are able to work with weak diffraction anisotropy. But severe diffraction anisotropy may represent a serious threat. The difficulties may appear in the process of phasing and/or structure refinement. However, several computational tools have been developed to analyse or even account for diffraction anisotropy, e.g. *AIMLESS* (Evans & Murshudov, 2013), *STARANISO* (Tickle *et al.*, 2018) and *Diffraction Anisotropy Server* (Strong *et al.*, 2006). These methods perform anisotropic cut-off of the data together with rescaling of intensities or structure factors with scales depending on the analysis and model of anisotropy employed by each program. These modifications are beneficial for a large number of crystal structures and are reported in Section 2 (Rupp, 2018).

Paired refinement is a modern method for determining the high-resolution cut-off of diffraction data (Karplus & Diederichs, 2012). For this method, the reference data are selected on the basis of a conservative cut-off [e.g. $\langle I/\sigma(I) \rangle < 2$]. More and more reflections are added to the model refinement in a stepwise manner, and their positive or negative contribution is evaluated on a number of criteria, mainly R_{free} calculated on the reference data. Recently, the method has been implemented in the program *PAIREF* (Malý *et al.*, 2020, 2021). However, the current protocol implemented in *PAIREF* does not consider the anisotropic diffraction qualities of the crystals. Both the reference data and the evaluated reflections are in the form of spherical shells.



OPEN ACCESS
Published under a CC BY 4.0 licence

We investigated the possibility of combining corrections for diffraction anisotropy with the standard paired refinement approach. The crystals of our target protein H33 showed serious anisotropy in the diffraction qualities. H33 is an artificial protein binder that was selected during a directed evolutionary study (Pham *et al.*, 2021); it is a variant of the protein scaffold derived from the N-terminal domain of the PIH1D1 domain of the R2TP cochaperone complex (PDB entry 4psf; Hořejší *et al.*, 2014). The scaffold was trained using the ribosome display technique to bind human interleukin-10 (IL-10), a cytokine of human innate immunity (El Kasmi *et al.*, 2007). Blocking or potentiating IL-10 signalization by artificially evolved non-antibody binders such as H33 could be an important component of the treatment of inflammatory, malignant and autoimmune diseases in which IL-10 plays a role. However, our understanding of the structural aspects of the binding between the binders and IL-10 is quite limited. So far, we have solved the structure of only one IL-10 binder called J61 (PDB entry 7avc; Pham *et al.*, 2021). Therefore, a newly solved structure of H33 will aid in the design of new potent and selective binders.

In our work, we introduced an approach to perform paired refinement using the anisotropic data. Anisotropic scaling proved to have a positive impact on the quality of the observed electron density.

2. Materials and methods

2.1. Protein production and crystallization

Protein production, purification and basic characterization have been described previously (Pham *et al.*, 2021). Briefly, the synthesized DNA strings were cloned into the pET-26b(+) vector. The plasmid was used for transformation into the *Escherichia coli* strain BL21(DE3). The bacteria were grown in LB medium, and protein expression was induced by the addition of isopropyl-beta-D-thiogalactopyranoside. After cell disruption, the soluble fraction was separated by centrifugation and the protein was purified from the cell lysate by affinity chromatography using Strep-Tactin XT resin. The last purification step was performed using size-exclusion chromatography (Superdex 75 16/600 column).

The crystals were prepared using the hanging-drop vapour-diffusion method from a protein solution that contained 20 mM Tris, 100 mM NaCl pH 8.0 and the protein at a concentration of 10 mg ml⁻¹. The protein crystallized in a wide range of crystallization conditions. However, the crystals diffracted poorly. The final crystallization conditions were 1 M (NH₄)₂SO₄, 1% (w/v) PEG 3350, 0.1 M bis-Tris pH 5.5. Cryo-protection with 20% (v/v) glycerol was necessary before flash-freezing in liquid nitrogen.

2.2. Diffraction data collection and processing

The synchrotron data were collected on beamline P13 (Cianci *et al.*, 2017) operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). The diffraction images were processed with *XDS* (Kabsch, 2010)

up to 2.3 Å resolution. The data quality metrics [decrease in $\langle I/\sigma(I) \rangle$, decrease in $CC_{1/2}$] indicate radiation damage that started immediately after 180° of total oscillation and progressed to the end of the measurement. Therefore, only half of the images (3600) were used for further data evaluation. Such data treatment should remove the possible impact of absorbed dose on the resulting diffraction anisotropy. Initial scaling of the data was performed using *AIMLESS* (Evans & Murshudov, 2013). The data were severely anisotropic according to a number of indicators. For example, the estimates of the diffraction limits reported by *AIMLESS* [based on criterion for $\langle I/\sigma(I) \rangle > 1.5$ in the highest-resolution shell] were 3.28 and 2.65 Å along the hk plane and the l axis, respectively.

Due to severe diffraction anisotropy, the unmerged scaled data from *XDS* (*XDS_ASCII.HKL* file) were merged and corrected for anisotropy using the *STARANISO* server (Tickle *et al.*, 2018) with four different local spherical ($\langle I_{\text{mean}}/\sigma(I_{\text{mean}}) \rangle$) cut-offs going down from 1.2 (*STARANISO* default value, A1.2 data) to 1.0 (A1.0 data), 0.75 (A0.75 data) and the lowest available value 0.5 (A0.5 data). The free flags were generated with the program *FREERFLAG* (Brünger, 1992). Initially, free flags were generated for the A1.2 data. The free flags for the A1.0 data were generated with the option to copy already existing flags for reflections in the A1.2 data. A similar approach was used for the generation of free flags for the A0.75 and A0.5 data. This approach was necessary to maintain the pairwise consistency of the free flags within the different data. Data quality indicators are shown in Table 1.

The phase problem was solved by molecular replacement using *PHASER* (McCoy *et al.*, 2007) employing the J61 variant of the protein binder from the same directed evolutionary study (PDB entry 7avc; Pham *et al.*, 2021) as a search model. Data with the highest $\langle I_{\text{mean}}/\sigma(I_{\text{mean}}) \rangle$ cutoff (A1.2) were used. Two molecules were found in the asymmetric unit. Due to the low resolution of the data and the unstable refinement (unacceptable number of Ramachandran outliers and bad bond angles) in *REFMAC5* (Murshudov *et al.*, 2011), the structure was restrained to the original scaffold of the PDB entry 4psf refined at 1.58 Å (Hořejší *et al.*, 2014) with *PROSMART* (Nicholls *et al.*, 2012). The structure was refined with isotropic atomic displacement parameters (ADPs) and no TLS domains defined. Manual corrections to the model were performed with *Coot* (Emsley *et al.*, 2010).

For the manually launched paired refinement, the data with the highest cut-off (A1.2) were initially chosen. Refinement of the structure model restrained to the structure of PDB entry 4psf with *REFMAC5* was used. To keep the same refinement scheme as used previously, three cycles were performed in each paired refinement step. We also performed several manual paired refinements with ten cycles of refinement. Although the results differ in exact values, this change did not lead to a different decision on data usage. Several criteria were evaluated during the paired refinement. Mainly, drops in overall R_{work} and R_{free} were monitored. In addition to that, R_{free} in the highest-resolution shell did not exceed the value of 0.42 (the theoretically perfect model gives an R value of 0.42

Table 1
Data collection and processing statistics.

Values in parentheses are for the highest-resolution isotropic shell.

Diffraction source	Petra III, P13				
Wavelength (Å)	0.976				
Temperature (K)	100				
Crystal-to-detector distance (mm)	576.1				
Exposure per image (s)	0.04				
Images collected/processed	7200/3600				
Space group	I4 ₂ 2				
<i>a</i> = <i>b</i> , <i>c</i> (Å)	123.13, 190.11				
Dataset	A1.2	A1.0	A0.75	A0.5	Iso
Local (<i>I</i> / σ (<i>I</i>)) cut-off	1.2	1.0	0.75	0.5	Isotropic data
Resolution range (Å)	47.53–3.29†–2.58 (2.83–2.58)	47.53–3.26†–2.56 (2.80–2.56)	47.53–3.20†–2.52 (2.76–2.52)	47.53–3.13†–2.47 (2.71–2.47)	47.53–2.80 (2.95–2.80)
Total No. of reflections	198331 (8000)	204983 (8055)	215735 (8090)	230863 (8194)	237010 (36042)
No. of unique reflections	15698 (785)	16246 (812)	17151 (858)	18374 (919)	18372 (2625)
Spherical completeness (%)	67.4 (14.4)	68.1 (14.6)	68.2 (14.3)	69.2 (14.8)	100 (100)
Ellipsoidal completeness (%)	94.5 (79.7)	94.6 (77.9)	94.3 (73.8)	94.2 (71.3)	N/A
Multiplicity	12.6 (10.2)	12.6 (9.9)	12.6 (9.4)	12.6 (8.9)	12.9 (13.7)
<i>I</i> / σ (<i>I</i>)	17.3 (1.9)	16.7 (1.6)	15.9 (1.3)	14.8 (0.8)	14.0 (1.0)
<i>R</i> _{meas}	0.114 (1.843)	0.117 (2.045)	0.121 (2.546)	0.127 (3.775)	0.129 (5.545)
<i>R</i> _{pin}	0.032 (0.559)	0.033 (0.629)	0.034 (0.801)	0.036 (1.216)	0.036 (1.476)
<i>CC</i> _{1/2}	0.999 (0.590)	0.999 (0.529)	0.999 (0.436)	0.999 (0.224)	0.999 (0.628)
<i>CC</i> *	0.999 (0.861)	0.999 (0.832)	0.999 (0.780)	0.999 (0.605)	0.999 (0.878)

† Lowest high-resolution diffraction limit after anisotropic cut-off.

against random data with no twinning and no translational non-crystallographic symmetry; Evans & Murshudov, 2013), and the values of *CC*_{work} and *CC*_{free} did not exceed the value of *CC** (see Table 1). The main results of the paired refinement are shown in Table 2. The decrease in *R*_{work} and *R*_{free} values in all three steps indicates that the addition of the progressively weaker reflections improved the model quality against the same (stronger) data. Therefore, A0.5 data were used in further structure refinements. The exact values of the final *R*_{work} and *R*_{free} in the fifth column of Table 2 cannot be directly compared because they were calculated against different data. Although the differences in the *R*_{work} and *R*_{free} values can be considered marginal, they are comparable to values published in previous studies (Karplus & Diederichs, 2012; Malý *et al.*, 2020, 2021).

The final model refinement using ten cycles was carried out using all reflections (work and free) of the A0.5 dataset. Jelly body protocol was used to release the previously used and necessary restraints. The quality of the structure stereochemistry was checked using the validation tools in *Coot* (Emsley *et al.*, 2010), *CCPA* (Agirre *et al.*, 2023; Winn *et al.*, 2011), *MOLPROBITY* (Chen *et al.*, 2010) and the Protein Data Bank (Berman *et al.*, 2003). The quality indicators of the final structure refinement are shown in Table 3. Raw diffraction data are available from <https://doi.org/10.5281/zenodo.4033811>. The structure coordinates were deposited under PDB entry 8bdu.

For analysis of the additional value of anisotropic scaling along with paired refinement in terms of data quality and observed electron density, data processed in the standard (isotropic) way were used in paired refinement with a 2.9 Å starting resolution. The complete cross-validation procedure implemented in *PAIREF* extended the resolution to 2.8 Å.

Table 2

Progress of paired refinement using data with a continuously decreasing (*I*/ σ (*I*)) cutoff.

Initial *R* values in *X*→*Y* steps are calculated using the model refined with data *X* against data *X*. Final *R* values are calculated using the model refined with data *Y* against data *X*.

Dataset	Resolution (Å)	Unique reflections added	Initial <i>R</i> _{work} / <i>R</i> _{free}	Final <i>R</i> _{work} / <i>R</i> _{free}	Δ <i>R</i> _{work} / <i>R</i> _{free}
A1.2	47.53–3.29† –2.58			0.2228/ 0.2560	
A1.2→A1.0	47.53–3.26† –2.56	548	0.2228/ 0.2560	0.2202/ 0.2525	–0.0026/ –0.0035
A1.0→A0.75	47.53–3.20† –2.52	905	0.2248/ 0.2572	0.2214/ 0.2538	–0.0034/ –0.0034
A0.75→A0.5	47.53–3.13† –2.47	1123	0.2287/ 0.2611	0.2247/ 0.2583	–0.0040/ –0.0028

† Lowest high-resolution diffraction limit after anisotropic cut-off.

Table 3

Structure refinement statistics.

Values in parentheses are for the highest isotropic resolution shell.

Resolution (Å)	47.53–3.13†–2.47 (2.71–2.47)
No. of residues/non-H atoms	266/2071
<i>R</i> _{work}	0.219 (0.355)
<i>R</i> _{free}	0.255 (0.368)
<i>CC</i> _{work}	0.993 (0.492)
<i>CC</i> _{free}	0.913 (0.590)
No. free reflections	910
<i>R</i> _{all}	0.223 (0.379)
Average ADP (Å ²)	82
R.m.s. deviations from ideal	
Bond lengths (Å)	0.010
Angles (°)	1.755
Ramachandran plot†	
Allowed	261 (99.6%)
Outliers	1 (0.4%)
Solvent content (%)	76

† As calculated by *MOLPROBITY*. † The lowest high-resolution diffraction limit after the anisotropic cut-off.

short communications

3. Results and discussion

The artificially generated binder H33 was successfully crystallized and the diffraction data were collected. The crystal diffracted anisotropically, and correction of the intensities for diffraction anisotropy was performed. The crystal structure was solved and refined using the A1.2 data. The high-resolution diffraction limit was extended using the paired refinement procedure to that of the A0.5 data. The A0.5 data were used for final structure refinement.

The structure of binder H33 is highly similar to that of binder J61 (Pham *et al.*, 2021) from the same study. The root mean square deviation calculated on 128 C α atoms is lower than 1.2 Å. The structure has an unusually high solvent content of 76% (Kantardjiev & Rupp, 2003). This high solvent

content is present in <2% of the crystal structures in the PDB. The solvent content is probably responsible for the low diffraction quality of the crystals.

Previous studies have shown that diffraction anisotropy is not strictly dependent on crystal packing (Robert *et al.*, 2017). The molecules in the crystal of binder H33 are arranged in tubules perpendicular to the *z* axis of the crystal lattice. The tubules have large channels of solvent between them. The planes with the normal vector perpendicular to the *z* axis are the least occupied with molecules [see Fig. 1(b)]. In contrast, no large channels are present in the planes with normal vectors perpendicular to the *x* or *y* axis.

Using the data range according to paired refinement may result in an improvement of the observed electron density (Karplus & Diederichs, 2015; Malý *et al.*, 2020). Correction for

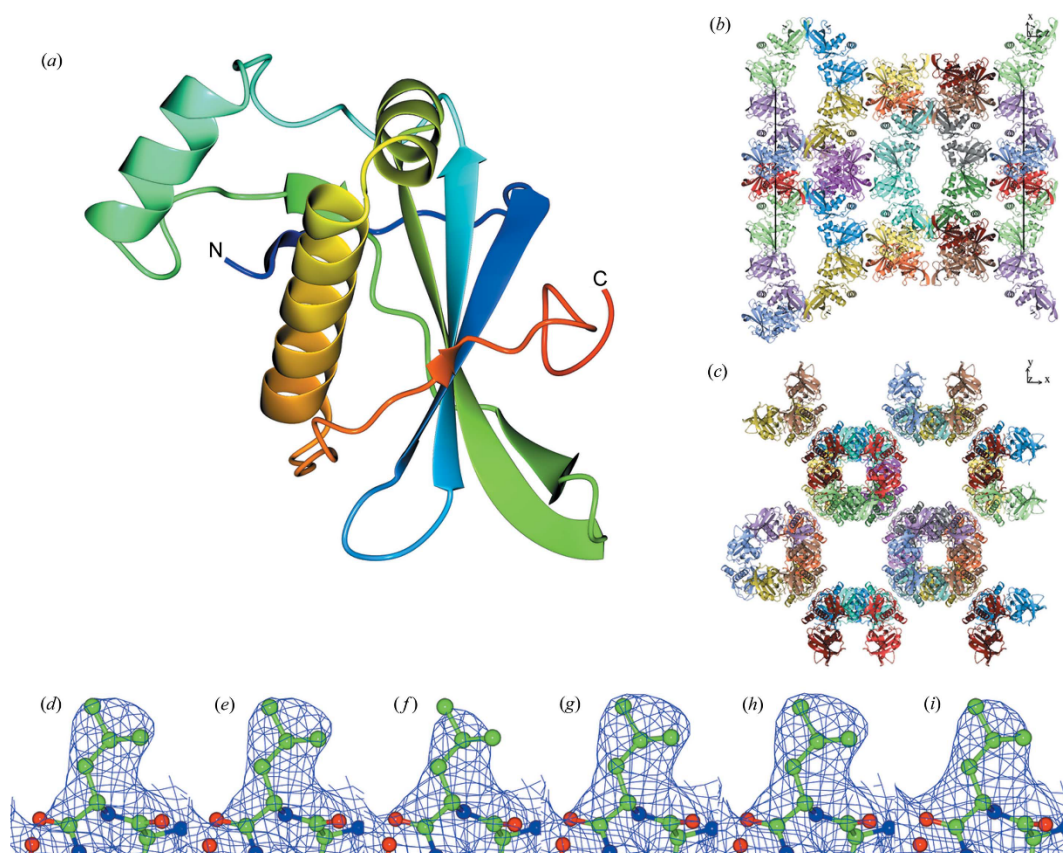


Figure 1

(a) Structure of the monomeric binder H33 determined using the A0.5 data in a secondary structure representation. (b)–(c) Unit cell filled with molecules viewed from different perspectives. (d)–(f) Residue Leu75 (chain A) with the $2mF_o - DF_o$ electron density (blue mesh) at the level of 0.15 e \AA^{-3} calculated for the model before the paired refinement procedure using data A1.2, the output model from the paired refinement procedure using the A0.5 data and the model from the isotropic paired refinement procedure using the Iso data, respectively. (g)–(i) Residue Leu75 (chain A) with the $2mF_o - DF_o$ electron density at the 1σ level for the same combination of model versus data as in the previous triplicate. Graphics were generated with CCP4MG (McNicholas *et al.*, 2011).

the anisotropy in the diffraction qualities was also shown to improve the observed electron density (Tickle *et al.*, 2018). Paired refinement using shells reflecting the diffraction anisotropy is not automated in any pipeline. The available software, for example *PAIREF* (Malý *et al.*, 2020) and *PDB-REDO* (Joosten *et al.*, 2014), use the addition of reflections in spherical shells by increasing the spherical high-resolution diffraction limit. Here, we propose the addition of reflections with the same expected information content in non-spherical shells.

The quality of the electron density depends on the diffraction data and the structural model. In our analysis, we compared electron density maps of (i) the starting model for the paired refinement refined using the A1.2 data against the A1.2 data, (ii) the resulting model from the paired refinement refined using the A0.5 data against the A0.5 data and (iii) the optimal model from the paired refinement using the Iso data at 2.8 Å resolution. The electron density maps were calculated with fast Fourier transformation using the same grid spacing to avoid possible bias (Urzhumtsev *et al.*, 2014). The electron density map from the Iso data is the least detailed [see Fig. 1(*f*)]. Corrections for diffraction anisotropy using the *STARANISO* server (Tickle *et al.*, 2018) and using the A1.2 data dramatically improved the quality of the observed electron density maps. No differences were observed with the extension of data from A1.2 to A0.5.

The number of reflections in the A0.5 data is approximately equal to that of the Iso data. Apparently, their information content is different. The anisotropic cut-off in the A0.5 data removed a significant portion of noisy reflections. The Iso dataset contains reflections in the weak directions with a low signal-to-noise ratio. Moreover, it does not contain a portion of the strong reflections in the strong directions that are present in the A0.5 data at a resolution higher than 2.8 Å.

In our case, both approaches to data optimization (correction for diffraction anisotropy and paired refinement) have proved useful. Although improvement in observed electron density did not occur after paired refinement of data corrected for diffraction anisotropy, 2676 unique reflections (14.5% from 18 374 reflections in total) were added to the refinement scheme using the A0.5 data. This addition was validated by the decrease in *R* values (see Table 2).

The current trend in data quality evaluation (paired refinement) is to investigate the 'additional value' of more and more observations involved in model refinement. Conventional indicators of the quality of the diffraction data are no longer relevant for the estimation of the high-resolution diffraction limit. The diffraction anisotropy makes the problem even more difficult. Our crystal structure was determined at a nominal diffraction limit of 2.47 Å. However, closer inspection of the diffraction data statistics shows that the diffraction data become dramatically incomplete at better than 3.13 Å resolution. The highest-resolution shell of reflections has a spherical completeness lower than 15%. This indicator must be considered when comparing structures refined 'with the same diffraction limits'.

Acknowledgements

We thank Isabel Bento for her assistance in using the beamline.

Funding information

This work was supported by the Ministry of Education, Youth and Sports CR, project CAAS (grant No. CZ.02.1.01/0.0/0.0/16_019/0000778 awarded to the Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague); the institutional grant to the Institute of Biotechnology of the Czech Academy of Sciences (RVO 86652036). We acknowledge CMS-Biocev (Crystallization and diffraction techniques) supported by MEYS CR (grant No. LM2018127), the Grant Agency of the Czech Technical University in Prague (grant No. SGS22/114/OHK4/2T/14).

References

- Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilstein-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M. & Yamashita, K. (2023). *Acta Cryst. D* **79**, 449–461.
- Berman, H., Henrick, K. & Nakamura, H. (2003). *Nat. Struct. Mol. Biol.* **10**, 980.
- Brünger, A. T. (1992). *Nature*, **355**, 472–475.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst. D* **66**, 12–21.
- Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S. & Schneider, T. R. (2017). *J. Synchrotron Rad.* **24**, 323–332.
- El Kasmí, K. C., Smith, A. M., Williams, L., Neale, G., Panopolous, A., Watowich, S. S., Häcker, H., Foxwell, B. M. & Murray, P. J. (2007). *J. Immunol.* **179**, 7215–7219.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D* **66**, 486–501.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst. D* **69**, 1204–1214.
- Hořejší, Z., Stach, L., Flower, T. G., Joshi, D., Flynn, H., Skehel, J. M., O'Reilly, N. J., Ogródowicz, R. W., Smerdon, S. J. & Boulton, S. J. (2014). *Cell Rep.* **7**, 19–26.
- Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCr*, **1**, 213–220.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Kantardjiev, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
- Karplus, P. A. & Diederichs, K. (2015). *Curr. Opin. Struct. Biol.* **34**, 60–68.

short communications

- Malý, M., Diederichs, K., Dohnálek, J. & Kolenko, P. (2020). *IUCrJ*, **7**, 681–692.
- Malý, M., Diederichs, K., Dohnálek, J. & Kolenko, P. (2021). *Acta Cryst. F77*, 226–229.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst. D67*, 386–394.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D67*, 355–367.
- Nicholls, R. A., Long, F. & Murshudov, G. N. (2012). *Acta Cryst. D68*, 404–417.
- Pham, P. N., Huličiak, M., Biedermannová, L., Černý, J., Charnavets, T., Fuertes, G., Herynek, Š., Kolářová, L., Kolenko, P., Pavlíček, J., Zahradník, J., Mikulecký, P. & Schneider, B. (2021). *Viruses*, **13**, 190.
- Robert, X., Kassis-Sahyoun, J., Ceres, N., Martin, J., Sawaya, M. R., Read, R. J., Gouet, P., Falson, P. & Chaptal, V. (2017). *Sci. Rep.* **7**, e17013.
- Rupp, B. (2018). *Structure*, **26**, 919–923.
- Strong, M., Sawaya, M. R., Wang, S., Phillips, M., Cascio, D. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
- Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vornrhein, C. & Bricogne, G. (2018). *STARANISO*. Global Phasing Ltd, Cambridge, UK. <https://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>.
- Urzhumtsev, A., Afonine, P. V., Lunin, V. Y., Terwilliger, T. C. & Adams, P. D. (2014). *Acta Cryst. D70*, 2593–2606.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst. D67*, 235–242.



ISSN 2059-7983

Crystal structure of human interferon- γ receptor 2 reveals the structural basis for receptor specificity

Pavel Mikulecký,† Jirí Zahradník,† Petr Kolenko, Jiří Černý, Tatsiana Charnavets, Lucie Kolářová, Iva Nečasová, Phuong Ngoc Pham and Bohdan Schneider*

Institute of Biotechnology CAS, BIOCEV, Prumyslova 595, 252 50 Vestec, Czech Republic. *Correspondence e-mail: bohdan.schneider@gmail.com

Received 17 May 2016

Accepted 27 July 2016

Edited by Z. S. Derewenda, University of Virginia, USA

† These authors share first authorship.

Keywords: interferon- γ receptor 2; fibronectin type III domain; class 2 cytokine receptors.

PDB reference: interferon- γ receptor 2, 5eh1

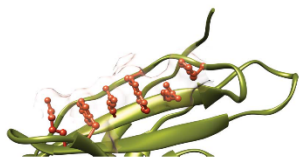
Supporting information: this article has supporting information at journals.iucr.org/d

Interferon- γ receptor 2 is a cell-surface receptor that is required for interferon- γ signalling and therefore plays a critical immunoregulatory role in innate and adaptive immunity against viral and also bacterial and protozoal infections. A crystal structure of the extracellular part of human interferon- γ receptor 2 (IFN γ R2) was solved by molecular replacement at 1.8 Å resolution. Similar to other class 2 receptors, IFN γ R2 has two fibronectin type III domains. The characteristic structural features of IFN γ R2 are concentrated in its N-terminal domain: an extensive π -cation motif of stacked residues KWRWRH, a NAG–W–NAG sandwich (where NAG stands for *N*-acetyl-D-glucosamine) and finally a helix formed by residues 78–85, which is unique among class 2 receptors. Mass spectrometry and mutational analyses showed the importance of N-linked glycosylation to the stability of the protein and confirmed the presence of two disulfide bonds. Structure-based bioinformatic analysis revealed independent evolutionary behaviour of both receptor domains and, together with multiple sequence alignment, identified putative binding sites for interferon- γ and receptor 1, the ligands of IFN γ R2.

1. Introduction

Interferon- γ receptor 2 is a cell-surface receptor that represents a crucial molecule in the interferon- γ (IFN γ) signalization pathway, influencing innate and adaptive immunity against pathogens and tumours (Schoenborn & Wilson, 2007; Lin & Young, 2013). The signalling cascade is initiated by the binding of IFN γ to its high-affinity cell surface receptor 1, forming a binary complex with a structure that has already been determined [PDB entries 1fg9 (Thiel *et al.*, 2000) and 1fyh (Landar *et al.*, 2000)]. However, to activate this binary complex and activate the JAK/STAT signalization pathway (Jung *et al.*, 1987; Cook *et al.*, 1994; Hemmi *et al.*, 1994), IFN γ receptor 2 must participate in a ternary complex created by a homodimer of IFN γ , two molecules of receptor 1 and molecule(s) of receptor 2. To date, detailed structural and biophysical characterization of IFN γ receptor 2, the key molecule for proper IFN γ signalization, is lacking.

From its sequence similarity, IFN γ receptor 2 (also known as IFN γ receptor β chain or accessory factor 1, AF-1), has been classified as a member of the class 2 receptor family. This large group of cytokine receptors includes IFN γ receptor 1, receptors of interferon- α and interferon- β , receptors of interleukin-10 and interleukin-20, and receptors of other interleukins belonging to the IL-10 family (Langer *et al.*, 2004). The mature IFN γ receptor 2 protein comprises of 310 amino acids and has a predicted molecular mass of 35 kDa. It consists of a relatively short 69-amino-acid intracellular domain, a 21-amino-acid transmembrane domain and a 220-



OPEN ACCESS

amino-acid extracellular domain that is structured into two fibronectin type III domains. The extracellular domain contains five cysteine residues and six potential N-linked glycosylation sites. Such extensive glycosylation contributes to a significant size heterogeneity, which is observed even when the receptor molecule is isolated from the same cell type; the molecular weight of mature receptor 2 of human interferon- γ ranges from 61 to 67 kDa (Bach *et al.*, 1995).

Despite its biological significance, three-dimensional structural data on IFN γ receptor 2 are lacking. Here, we report a 1.8 Å resolution crystal structure of the extracellular portion of IFN γ receptor 2 (hereafter called IFN γ R2); the structure has been deposited in the Protein Data Bank as entry 5eh1. The structure and sequence of IFN γ R2 are discussed in the context of the structures and sequences of the related class 2 cytokine receptors, with emphasis on the sequentially closest receptors of interleukins from the IL-10 family. Structure-based and sequence-based alignments suggested regions securing binding specificity of these receptors for their cytokine ligands.

2. Materials and methods

2.1. Cloning, expression and purification of IFN γ R2

The gene encoding the extracellular part of IFN γ R2 (residues 28–247 of UniProt entry P38484) was cloned into a *Drosophila* pMT-BiP-V5-His_A vector using BglIII and AgeI restriction enzymes in frame with an N-terminal BiP signal peptide and a C-terminal 6 \times His tag. This expression vector was co-transfected into insect Schneider S2 cells along with the pCoBlast selection plasmid using Effectene Transfection Reagent according to the manufacturer's instructions. Blasticidin-resistant S2 cells were selected by growing the cells in HyClone SFX Medium supplemented with 10% FBS and 25 μ g ml⁻¹ blasticidin S. Large-scale protein expression was achieved after expansion and substitution into HyClone SFX serum-free medium, and protein expression was induced by the addition of 0.75 mM CuSO₄ for 6 d (the cell concentration was approximately 35 million per millilitre) and 1.5 mM CuSO₄ for a further 2 d until the percentage of living cells did not decrease below 95%. After expression, the cells were discarded by centrifugation and the medium containing secreted glycosylated IFN γ R2 protein was supplemented with the following additives at the following final concentrations: 5 mM CaCl₂, 1 mM NiSO₄, 250 mM NaCl and 50 mM Tris-HCl pH 8. The protein was purified on an IMAC HP column charged with NiSO₄ and equilibrated with EQ buffer (50 mM Tris-HCl pH 8, 500 mM NaCl). The column was washed with W buffer (50 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM imidazole pH 8) and the protein was eluted with EL buffer (50 mM Tris-HCl pH 8, 500 mM NaCl, 250 mM imidazole pH 8). It was further purified to homogeneity by size-exclusion chromatography at room temperature on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) equilibrated with HN buffer (10 mM HEPES pH 7.5, 100 mM NaCl). Samples were analyzed by 12% SDS-PAGE.

IFN γ R2 was produced in insect cells as a secreted protein bearing oligosaccharide moieties of approximately 10 kDa according to SDS-PAGE analysis. Deglycosylation by peptide: N-glycosidase F (PNGase F) or endoglycosidase H (Endo H) with a C-terminal *Strep*-tag (§S1, Supporting Information) was performed after purification of IFN γ R2 on an IMAC column during dialysis against TN buffer (50 mM Tris buffer pH 8, 150 mM NaCl) or HN buffer, respectively. Endoglycosidases were removed on a *Strep*-Tactin column and the nonbound fraction containing IFN γ R2 was further purified by size-exclusion chromatography in HN buffer.

The single IFN γ R2 variants (N110Q, N137Q and N231Q, respectively) were introduced using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Primers are listed in Supplementary Table S1. The fully mutated IFN γ R2 variant bearing N56Q, N110Q, N137Q and N231Q mutations was obtained as a GeneArt Strings DNA Fragment and was cloned with the same protocol as the wild type. The expression and purification of all IFN γ R2 variants were performed in the same way as described above.

2.2. Biophysical measurements

Circular-dichroism (CD) spectra were recorded using a Chirascan-plus spectrometer (Applied Photophysics) in steps of 1 nm over the wavelength range 185–260 nm. Samples diluted with water to a concentration of 0.2 mg ml⁻¹ were placed into the holder in a 0.05 cm path-length quartz cell and individual spectra were recorded at a temperature of 23°C. The CD signal was expressed as the ellipticity and the resulting spectra were buffer-subtracted. To analyze the ratio of secondary structures, we used the *CDNN* program (Böhm *et al.*, 1992) provided with the Chirascan CD spectrometer. CD melting measurements were performed using samples diluted with water to a protein concentration of 0.5 mg ml⁻¹. A 10 mm path-length quartz cell was placed into the thermostated holder and sample absorption was recorded at 280 nm in 1°C increments at a rate of 0.5°C min⁻¹ over the temperature range 20–85°C with an averaging time of 12 s. Melting curves were normalized to relative values between 0.0 and 1.0 to visually magnify differences between the melting profiles, and the melting temperature (T_m) was estimated from the first derivative of the melting curves.

2.3. Glycosylation analysis and disulfide-bond determination

IFN γ R2 glycosylation sites were determined by MALDI-MS analysis preceded by protein digestion as described previously (Plihal *et al.*, 2004). Disulfide bonds in IFN γ R2 were determined by SDS-PAGE and subsequent identification by mass spectrometry (MS) in analogy to the previously described procedure (Pompach *et al.*, 2009). 20 μ g of sample in nonreducing conditions was loaded onto a 4–12% gradient gel (Life Technologies) in the presence of 200 μ M cystamine. Bands corresponding to highly glycosylated IFN γ R2 were excised and subjected to in-gel deglycosylation and proteolysis. Deglycosylation using Endo H (New England Biolabs) was carried out for 4 h at 37°C and the resulting

Table 1
Data-collection statistics and structure-refinement parameters.

Values in parentheses are for the highest resolution shell.	
X-ray source	MX 14.1, HZB
Wavelength (Å)	0.91841
Total oscillation angle (°)	180
Resolution range (Å)	62.88–1.80 (1.91–1.80)
Space group	<i>P</i> 6 ₂ 2
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 58.102, <i>c</i> = 377.266
No. of measured reflections	688675 (110540)
No. of unique reflections	36723 (5733)
Average multiplicity	18.8 (19.3)
Completeness (%)	99.9 (99.8)
Average <i>I</i> σ(<i>I</i>)	17.6 (2.5)
Overall <i>B</i> factor from Wilson plot (Å ²)	20
Average <i>B</i> factor (Å ²)	28
No. of non-H atoms	
Protein	1734
Saccharides	42
Waters	311
All	2121
<i>R</i> _{merge}	0.148 (1.395)
Half-data-set correlation coefficient <i>CC</i> _{1/2}	99.9 (81.2)
No. of reflections, test set	1820
Final <i>R</i> _{work} / <i>R</i> _{free} / <i>R</i> _{all}	0.190/0.222/0.191
Ramachandran plot	
Residues in favoured region	213 [96.3%]
Residues in allowed regions	219 [99.1%]
Outliers	2 [0.9%]

partly deglycosylated sample was digested with trypsin (sequencing grade, Promega) for 12 h at 37°C at a protein:enzyme ratio of 30:1(w/w). After digestion, the tryptic peptide mixture was desalted on a peptide MicroTrap column (Michrom Bioresources) and separated on a reversed-phase C18 column (Acclaim PepMap 100, 5 µm, 0.1 × 20 mm; Thermo Scientific). The mobile phases consisted of 0.1% formic acid in 2% acetonitrile (solvent *A*) and 0.1% formic acid in 98% acetonitrile (solvent *B*). Peptides were eluted under the following gradient conditions: 2–45% solvent *B* in 40 min, 45–95% solvent *B* in 5 min. The flow rate was 0.5 µl min⁻¹ and the column was directly connected to the mass spectrometer. Mass spectra were acquired on a solarix XR FTMS instrument equipped with a 12 T superconducting magnet (Bruker Daltonics). For the identification of disulfide bonds, we used the *Links* algorithm, previously described as *ASAP* (*Automated Spectrum Assignment Program*; Schilling *et al.*, 2003). To generate deconvoluted spectra and export the *m/z* values, we used a script utilizing the *SNAP* 2.0 algorithm of the *DataAnalysis* 4.2 software suite (Bruker Daltonics).

2.4. Crystallization and diffraction data collection

Crystals of Endo H-deglycosylated IFNγR2 receptor were grown using the sitting-drop vapour-diffusion method in 96-3 three-well Intelli-Plate trays (Art Robbins Instruments). The reservoir solution consisted of 0.1 M MES pH 5.0, 10% PEG 6000 (final pH 6.0): condition No. 61 of The JCSG Core I Suite (Qiagen). Drops consisting of 0.2 µl protein sample (15 mg ml⁻¹ protein in HN buffer) and 0.2 µl reservoir solution were prepared with a Gryphon liquid-pipette robot (Dunn Labortechnik) and were equilibrated against 100 µl reservoir solution. Crystals appeared after 30 d of incubation

at 291 K. Crystals were mounted in Round LithoLoops (Molecular Dimensions) and flash-cooled in liquid nitrogen after cryoprotection in 20% (v/v) glycerol. X-ray diffraction data were collected at 100 K on beamline MX 14.1 of the BESSY II synchrotron-radiation source at the Helmholtz-Zentrum Berlin (HZB). A native data set was collected at a wavelength of 0.918 Å.

2.5. Data processing, structure determination and refinement

The diffraction and refinement statistics are summarized in Table 1. Diffraction data were processed and scaled using the *XDS* program package (Kabsch, 2010). The structure was solved with *BALBES* (Long *et al.*, 2008), but the structure model needed significant manual remodelling. Only the C-terminal domain of IFNγR2 was found and the initial *R* factors were about 0.49 and 0.52 for *R*_{work} and *R*_{free}, respectively. Residues missing from the initial model were built in with significant help from *ARP/wARP* (Langer *et al.*, 2008); manual corrections and building were performed using *Coot* (Emsley & Cowtan, 2004). Refinement was then carried out with *REFMAC5* (Murshudov *et al.*, 2011) and the structure was validated by *MolProbity* (Chen *et al.*, 2010). The coordinates and structure factors have been deposited in the PDB with accession code 5eh1.

2.6. Sequence and structural bioinformatics

The UniProt database was searched with the *BLAST* tool (Camacho *et al.*, 2009) using the sequence of the extracellular part of IFNγR2 as the query sequence. The automated result was manually reviewed to select 90 sequences from different species. These sequences were used to calculate a multiple sequence alignment with *Clustal Omega* (Sievers *et al.*, 2011) as implemented in *UGENE* (Okonechnikov *et al.*, 2012). The *ConSurf* server (Glaser *et al.*, 2003) was used to estimate the evolutionary conservation of amino-acid positions in the protein structures. The calculations were based on the crystal structure of IFNγR2 (PDB entry 5eh1) and the alignment prepared by *Clustal Omega*. Structural comparison was prepared by *MatchMaker* as implemented in the *UCSF Chimera* software (Pettersen *et al.*, 2004).

Root-mean-square deviation (r.m.s.d.) values between the N- and C-terminal domains were calculated using *VMD* (Humphrey *et al.*, 1996). The backbone atoms of 34 sequentially conserved residues in each domain (listed in Supplementary Table S2) were used for the structure superposition of all possible pairs of N- and C-terminal domains of the 12 available crystal structures of class 2 cytokine receptors. The *VMD* commands *measure fit* and *move* were used for the structural overlay, followed by *measure rmsd* operating on the same selection of residues and backbone atoms for the calculation of r.m.s.d. values. The stability of the N-terminal domain of IFNγR2 was estimated by calculating the pairwise interaction energy at the DFT-D level as detailed in Supplementary Fig. S3.

3. Results and discussion

3.1. Summary

IFN γ R2 was produced in insect S2 cells, purified, characterized by biophysical techniques (Fig. 1) and its crystal structure was solved at 1.8 Å resolution. Fig. 2 highlights some of the structural features of IFN γ R2; Figs. 3 and 4 provide a comparison of IFN γ R2 to the other cytokine receptors from the class 2 family, with the aim of correlating the sequences and structures of these proteins.

3.2. Glycosylation and overall fold stability

The IFN γ R2 protein has six potential glycosylation sites (Asn56, Asn110, Asn137, Asn231, Asn85 and Asn219), of which the first four were confirmed by mass-spectrometric analysis as glycosylated in our construct; the crystal structure later revealed that position Asn85 was also glycosylated. The greatest heterogeneity was observed at position Asn137. To remove oligosaccharide moieties from these residues, we used two endoglycosidases, Endo H and PNGase F, but both enzymes left several forms of IFN γ R2 with residual glycosylation as observed by a distribution of molecular mass on SDS-PAGE (Supplementary Fig. S1). Mass spectrometry identified α (1-3)-fucose at position Asn231. Because α (1-3)-fucose abolishes the activity of both endoglycosidases, its presence is a likely to be reason for the mass distribution of IFN γ R2. Deglycosylation by Endo H caused an approximately 7 kDa shift in molecular weight on SDS-PAGE and, in contrast to deglycosylation by PNGase, did not induce protein oligomerization, as checked by size-exclusion chromatography. Measurements using CD spectroscopy (Fig. 1) and thermal shift assay (§S1 in Supporting Information and Supplementary Fig. S2) showed no significant difference in melting temperatures between the glycosylated and Endo H-deglycosylated forms of the IFN γ R2 protein. Because the CD spectra of these two forms are also virtually identical, we

believe that deglycosylation does not influence the secondary structure of IFN γ R2. Although the CD spectra of IFN γ R2 and IFN γ R1 differ considerably (Černý *et al.*, 2015), both proteins belong to the same fold of the fibronectin type III domain family (Pfam PF00041).

Besides deglycosylation by the endoglycosidases, we designated asparagine-to-glutamine mutants to decrease the level of glycosylation. We prepared a fully mutated IFN γ R2 variant bearing N56Q, N110Q, N137Q and N231Q mutations and single-point mutants N110Q, N137Q and N231Q, respectively. All of these constructs were transfected into insect S2 cells, but none of them were secreted into the culture medium. This correlates with the earlier observation that IFN γ R2 mutants with changed glycosylation patterns were located in the cytoplasmic fraction (Moncada-Vélez, 2013). Structurally significant is glycosylation at positions Asn110 and Asn137, where the bound *N*-acetyl-D-glucosamine (NAG) residues sandwich Trp131 (Fig. 2b), thus shielding its hydrophobic surfaces from solvent. As suggested by the failure to express and/or purify the N56Q, N110Q, N137Q and N231Q mutants, glycosylation is necessary for IFN γ R2 production by stabilizing the fold and transport to and/or across the cellular membrane.

3.3. The IFN γ R2 fold is stable without disulfide bonds

IFN γ R2 contains five cysteine residues, and our mass-spectrometric analysis identified disulfide bonds linking Cys86 to Cys94 and Cys209 to Cys234. We observed the same protein mobility under nonreducing and reducing conditions during SDS-PAGE analysis. The melting temperature of both glycosylated and deglycosylated IFN γ R2 measured by thermal shift assay (Supplementary Fig. S2) decreased by only $\sim 1^\circ\text{C}$ in the presence of 5 mM TCEP (a reducing agent to break the disulfide bonds), so that the IFN γ R2 fold is stable without S-S bonds. This contrasts with the behaviour of IFN γ R1 (Fountoulakis *et al.*, 1990), in which the protein fold is

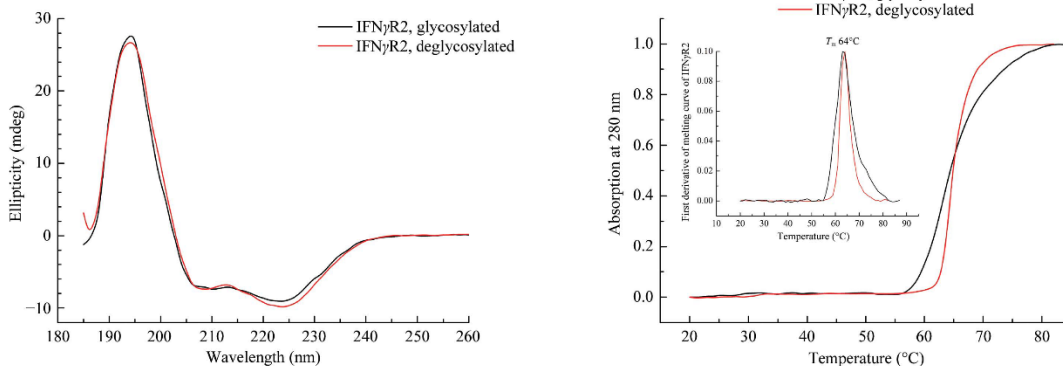


Figure 1 Left: circular-dichroism (CD) spectra of glycosylated and deglycosylated IFN γ R2. The CD spectra of both proteins are highly similar, suggesting that the partial removal of the oligosaccharide moieties did not affect the overall structure of IFN γ R2. Right: normalized melting curves measured from temperature-dependent CD spectra at 280 nm. The melting temperature was estimated as 64°C for both glycosylated and deglycosylated IFN γ R2.

stabilized to a large extent by S—S bridges. The fifth IFN γ R2 cysteine residue, Cys174, does not form an intramolecular S—S bridge but is bound to a monomeric cysteine. Binding of free cysteine to the sterically accessible Cys174 probably occurs after secreting IFN γ R2 into the cell-culture medium, which contains free cysteine and stabilizes the monomeric form of IFN γ R2.

3.4. The overall structure of the extracellular portion of IFN γ R2

The structure of IFN γ R2 was solved at 1.8 Å resolution and electron density was observed for amino-acid residues 28–240 of UniProt entry P38484, except for two two-residue loops. Data-collection and refinement parameters are shown in Table 1. The extracellular part of the IFN γ R2 molecule consists of two domains (Fig. 2), the N-terminal D1 domain of UniProt residues 28–133 and the C-terminal D2 domain of residues 144–247. Both domains belong to the immuno-

globulin fold with fibronectin type III topology, forming β -sandwiches (Pfam PF00041). The inter-domain torsion angle is approximately 120°, similar to those of IFN γ R1 (Thiel *et al.*, 2000; Walter *et al.*, 1995) and human tissue factor (Harlos *et al.*, 1994); the D1–D2 torsion angle is defined in Supplementary Table S3. The D1 domain is composed of three β -strands stacked on a layer of four β -strands, and the D2 domain is created by four β -strands arranged against four other β -strands; both domains are connected by a short linker (residues 134–143 in IFN γ R2) comprising a short helix that is also found in IFN γ R1, human tissue factor and other receptors.

3.5. Structural motifs in D1 and D2

D1 contains a distinct structural motif of six stacked residues: Lys68, Trp74, Arg114, Trp126, Arg116 and His123. The average distance between the mean planes of the individual side chains of this extensive π -cation interaction is 3.65 Å.

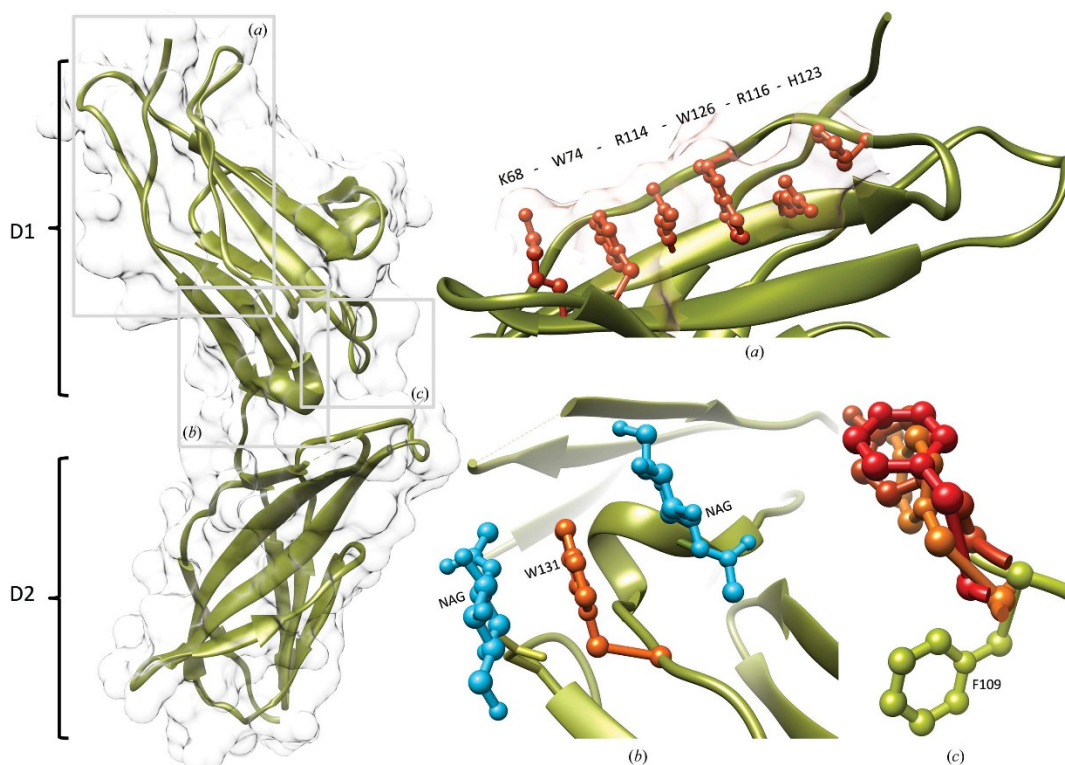


Figure 2
Left: ribbon and surface representations of the IFN γ R2 structure. D1 and D2 indicate domains 1 and 2, respectively. Insets: (a) residues Lys68, Trp74, Arg114, Trp126, Arg116 and His123 of the D1 domain form a stacking motif on the IFN γ R2 surface. (b) *N*-Acetyl-D-glucosamines (NAGs; blue) glycosylating Asn110 and Asn137 sandwich Trp131 (orange), reducing its hydrophobic character. (c) The superposition of aromatic binding epitopes shows differences between IFN γ R2 (Phe109 in green) on one side and promiscuous shared cytokine receptors on the other [in red; Tyr82 of IL10R2 (PDB entry 3lqm; Yoon *et al.*, 2010), Phe169 of gp130 (PDB entry 1bqj; Bravo *et al.*, 1998) and Tyr103 of γ_c receptor (PDB entry 4gs7; Ring *et al.*, 2012)]. No corresponding aromatic residue is observed in IL20R2 (Logsdon *et al.*, 2012).

research papers

Analysis of the interaction energies in D1 revealed that the motif contributes significantly to the overall stability of the whole domain. These six surface residues are involved in interactions that are comparable in strength to the hydrophobic core of the domain and are likely to play an important role in the process of domain folding. The residues responsible for domain stability are depicted in Supplementary Fig. S3, in which the colour and thickness of the cartoon representation show the relative interaction energy per residue ranging from low (blue) to high (red) stabilizing values.

An analogous stacking motif with the consensus sequence WSXWS (Bazan, 1990) has been predicted by sequence alignments in D1 of the class 1 receptor family (McElroy *et al.*, 2009), but such a motif is missing in the D2 domains of both class 1 and class 2 receptors. Based on the presence of the KWRWRH motif in IFN γ R2, we performed structural alignment of the class 2 receptor structures and discovered a similar but sequentially noncontinuous motif with the sequence (X)WRWR(X), where X is K, R or H. The important role of large aromatic tryptophan residues in stabilizing the fibronectin fold by stitching together two β -strands is accompanied in D1 by a structural role for charged residues, especially arginines. Besides the discussed (X)WRWR(X) motif, we found a tight overlap of a continuous chain of residues R-L/V-R-A (residues Arg114-Leu115-Arg116-Ala117 in IFN γ R2); the average r.m.s.d. between motifs from two receptors is 0.6 Å. The other important residues that are conserved in D1 of the available class 2 receptor structures are

residues corresponding to Trp49, Ser124 and the Cys86/Cys96 pair forming a disulfide bond in IFN γ R2. A unique feature of IFN γ R2 D1 is a short helix (residues 78–85), which is present in no other discussed receptor structure.

Sequential and structural comparison of D2 revealed a considerable sequence variability, within which we identified the conservation of two proline residues, Pro142 and Pro143, and the structurally well conserved motif 175-YNVA Δ XW-180, with r.m.s.d. values about 1 Å but low sequence similarity. Another characteristic structural feature of D2 is the formation of a disulfide bridge between Cys209 and Cys234. Higher values for the B factors in D2 indicate its higher flexibility compared with D1. A higher flexibility of D2 compared with D1 was also indicated in our previous studies of IFN γ R1 (Mikulecký *et al.*, 2013; Černý *et al.*, 2015).

3.6. Structural alignment of domains D1 and D2 in IFN γ R2 and in other class 2 receptors

We performed alignment of the IFN γ R2 structure with the 11 remaining available structures of the class 2 receptor family in order to gauge their similarity and reveal their unique features. The alignment was measured by overlapping 34 residues in the N-terminal D1 and the same number of residues in the C-terminal D2; the r.m.s.d. values of the overlapped residues are listed in Fig. 3 and the overlapped residues are listed in Supplementary Table S2. The D1 domains are mutually more similar than the D2 domains, as highlighted by

PDB code	1fg9	5eh1	1j7v	3lqm	4doh	4doh	3dlq	3g9v	3og6	3se4	3se4	2puq
Chain ID	C	A	R	A	E	B	R	A	B	A	C	T
Receptor	IFN γ R1	IFN γ R2	IL10R1	IL10R2	IL20R1	IL20R2	IL22R	IL22BP	IL28R	IFN α R1	IFN α R2	TF
IFN γ R1	0.82	0.87	0.73	0.73	0.80	0.78	0.90	0.71	0.82	1.35	1.12	0.85
IFN γ R2	1.27	0.79	0.93	0.97	0.79	0.79	0.96	0.80	0.85	1.45	1.25	1.04
IL10R1	1.22	1.12	1.07	0.74	0.97	0.88	0.94	1.43	1.13	0.81	0.67	0.66
IL10R2	1.00	1.14	1.05	0.89	0.95	1.05	0.65	0.87	0.68	0.80	1.11	0.93
IL20R1	0.93	1.32	0.89	1.33	1.59	0.93	0.97	0.63	0.75	0.59	1.36	1.15
IL20R2	1.32	1.51	1.23	0.99	1.33	0.91	0.93	0.66	0.82	1.48	1.25	1.21
IL22R	1.20	1.49	1.09	0.86	1.04	1.16	4.19	1.20	1.19	0.87	0.98	1.03
IL22BP	1.32	1.48	0.94	1.20	1.18	1.30	1.17	0.93	1.02	1.16	1.11	0.96
IL28R	1.21	1.16	1.34	1.29	1.30	1.32	1.18	1.45	0.91	1.09	0.77	0.86
IFN α R1	1.16	0.96	1.31	1.33	1.33	0.98	1.75	1.26	1.59	1.77	0.94	1.20
IFN α R2	1.35	1.80	1.40	1.29	1.19	1.81	1.58	1.14	1.46	1.22	1.43	1.02
TF	1.38	1.16	1.23	1.26	1.44	1.04	1.16	1.40	1.43	1.16	1.81	0.86

Figure 3
Structural differences between the N-terminal (D1) and C-terminal (D2) domains of 12 class 2 cytokine receptors gauged by the r.m.s.d. values for backbone atoms of their 34 residues. R.m.s.d. values comparing D1 and D2 domains are shown above and below the diagonal, respectively. For instance, comparison between D1 of IL20R2 and IL22BP gives an r.m.s.d. of 0.66 Å; the r.m.s.d. between their D2 domains is 1.30 Å. R.m.s.d. values that are smaller and larger than the off-diagonal average r.m.s.d. value are highlighted in blue and red hues, respectively. The diagonal (in grey) shows the lowest r.m.s.d. values for 34 residues from D1 and D2 within each receptor structure; the r.m.s.d. between D1 and D2 of IL10R2 is 0.89 Å. References to the analyzed structures are as follows: IFN γ R1, Thiel *et al.* (2000); IFN γ R2, this work; IL10R1, Josephson *et al.* (2001); IL10R2, Yoon *et al.* (2010); IL20R1 and IL20R2, Logsdon *et al.* (2012); IL22R, Bleicher *et al.* (2008); IL22BP, de Moura *et al.* (2009); IL28R, Miknis *et al.* (2010); IFN α R1 and IFN α R2, Thomas *et al.* (2011); TF (human tissue factor), Larsen *et al.* (2007).

blue and red hues in Fig. 3; the average r.m.s.d. between two D1 domains is 0.95 Å and that between two D2 domains is 1.3 Å. A high similarity within D1 and D2, respectively, indicates that modulation of the specificity of receptors takes place in only a few variable regions, which are discussed below.

D1 domains bear two conflicting structural features: strict fold conservation reflected by high structural similarity of the selected residues, and at the same time the presence of two structurally highly variable loops corresponding to residues 70–73 and 97–107 in IFN γ R2. The third variable loop was located in the D2 domain (residues 162–171 in IFN γ R2); the loops are coloured red and yellow for D1 and green for D2 in Fig. 4. Given the fairly uniform core of both domains and variability concentrated in the three localized regions, we suggest that the binding specificity of the individual receptors is controlled by these variable regions. However, there is another factor that contributes to the receptor specificity, the different mutual orientation of D1 and D2 (Supplementary Table S3), which displaces these variable regions to different positions, thus providing a unique binding interface for each receptor.

Several structural and sequential features of receptors of interferon- α and interferon- β (PDB entry 3se4; Thomas *et al.*, 2011), here labelled IFN α R1 and IFN α R2, distinguishes them from the other analyzed receptors. Specifically, IFN α R1 is composed of four instead of two domains; here, we analyzed D1 and D2. The D3–D4 pair cannot be analyzed as D4 is not resolved in the electron density. Further dissimilarities of IFN α R1 and IFN α R2 are found in the composition of their π -cation motifs: in IFN α R1 only four residues stack in D1 (Trp46, Arg76, Trp87 and Arg78) and three in D3 (Trp250, Arg279 and Trp291), while D1 of IFN α R2 does not have the motif at all and is replaced by the motif YVTV.

3.7. Similarity among receptors and consequences for evolution

As discovered previously (Yoon *et al.*, 2010), an aromatic tyrosine or phenylalanine residue situated in the cleft between D1 and D2 of gp130 (PDB entry 1bqu; Bravo *et al.*, 1998), γ_c (PDB entry 4gs7; Ring *et al.*, 2012) and IL10R2 (PDB entry 3lqm; Yoon *et al.*, 2010) serves as the key binding epitope of promiscuous class 1 and 2 receptors, implying the existence of a common ancestor. Superposition of these receptor structures shows that the orientation of Phe109 and three residues in gp130, γ_c and IL10R2 are quite different (Fig. 2c). No preferred rotamer of Phe109 overlaps the three former residues without significant rebuilding of the IFN γ R2 backbone. The structural difference between IFN γ R2 and the other receptors, especially IL10R2, is significant and suggests that there is not a common binding epitope for these receptors.

Significant sequence similarity between IFN γ R2, IL10R2 and IL20R2 (20–25% sequence identity among different species; analysis not shown) points to their evolutionary relationship. If proven, it would be analogous to the evolutionary relationship between receptors 1 of the cytokines IFN γ R1, IL10R1 and IL20R1 (Langer *et al.*, 2004). We may

therefore hypothesize that these three cytokine systems have evolved from a common ancestral system: while interferon- γ evolved early in evolution and is known in fish species (Savan *et al.*, 2009), its receptor 2 emerged later in connection with the evolution of amphibians. The specific function of IFN γ R2 therefore evolved from an older promiscuously functioning molecule. A likely candidate is IL10R2, because it is evolutionarily older and is known in primitive fishes, while IL20R2 emerges similarly to IFN γ R2 in amphibians. The lack of a common binding epitope between IFN γ R2 and the other class 2 receptors, notably IL10R2 (Fig. 2c), indirectly supports this hypothesis.

3.8. The sequence alignment of IFN γ R2 from various species suggests its binding interface

In an attempt to identify the putative interface by which IFN γ R2 forms a functional ternary complex with its binding partners interferon- γ and receptor 1, we aligned the IFN γ R2 sequences from 90 species and used the *ConSurf* server to project the consensus onto the structure of IFN γ R2 (Fig. 4b). The 32 conserved residues (purple in Fig. 4b) are mainly located in the inward arched part of the U-shaped receptor molecule. This part of the molecule contains the previously described stacking motif (Fig. 2a) and plays an important role in maintaining the overall structure. The 34 most variable residues are predominantly on the opposite side of the molecule (cyan in Fig. 4b). These regions of sequentially least conserved residues coincide with the location of the structurally variable loops derived by the superposition of receptor structures. We therefore conclude that the putative IFN γ R2 interface for forming the active ternary complex with IFN γ and IFN γ R1 is likely to be in the receptor 2 region with the most variable residues. This conclusion is supported by an analogous observation in the IFN γ R1 system: the least sequentially conserved residues form the interface with the binding partner (Mikulecký *et al.*, 2013).

The composition of the ternary signalling complexes of dimeric cytokines discussed here, IFN γ and IL-10, is understood less than for monomeric examples such as IL-20, for which the ternary complex has a known crystal structure (Logsdon *et al.*, 2012). One of the reasons is the existence of two binding interfaces in the dimeric cytokines and the resulting different and more complex stoichiometry of the complexes; the crystal structure of an IFN γ –IFN γ R1 complex with an unexpected 2:3 stoichiometry serves as an example (Thiel *et al.*, 2000). The topology and structure of the signalling ternary complex of IFN γ have been extensively studied and reviewed (Pestka *et al.*, 1997; Hoffmann *et al.*, 2015). Experiments in solution and on the cell surfaces indicated a 2:2:2 or 2:2:1 stoichiometry of the signalling IFN γ complex (Masters *et al.*, 1995); cross-linking of different components of the IFN γ complex expressed in cloned cell lines have shown direct contact between IFN γ R1 and IFN γ R2 (Krause *et al.*, 2006) and also between IFN γ and IFN γ R2 (Kotenko *et al.*, 1995). The newly determined structure of IFN γ R2 may spur new experiments exploring the topology and three-dimensional structure of the signaling ternary complex of IFN γ .

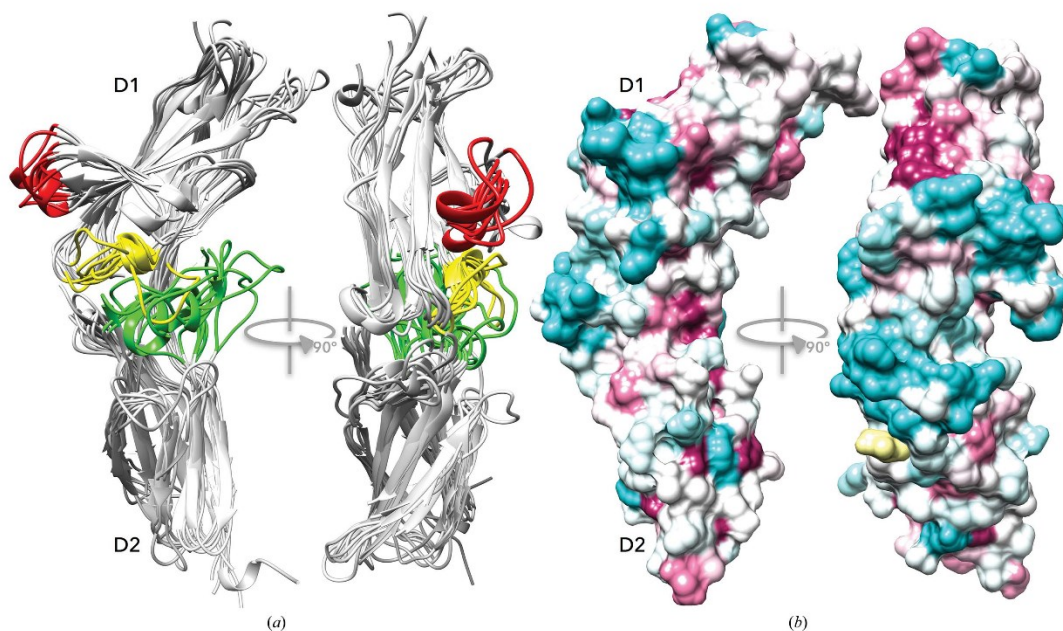


Figure 4
Structurally variable regions and a proposed interaction interface of the cytokine receptors. (a) Cartoon representation of the structural superposition of the D1 and D2 domains of the nine receptor structures. The backbone of the selected residues (listed in Supplementary Table S2) was superimposed independently for each domain and was then drawn on top of the IFN γ R2 structure. The aligned cores of both domains overlap tightly (for r.m.s.d. values, see Fig. 3), but two regions in D1 (red and yellow) and one in D2 (green) are highly variable. (b) Sequences of IFN γ R2 from 90 species with the sequentially most variable regions coloured cyan and conserved regions in purple drawn on the surface of the IFN γ R2 structure by *ConSurf* (Glaser *et al.*, 2003). The IFN γ R2 free cysteine Cys174 is highlighted in yellow. (a) and (b) show IFN γ R2 in the same orientation.

4. Conclusions

A partially deglycosylated extracellular part of the interferon- γ receptor 2, IFN γ R2, was crystallized and its structure was determined at 1.8 Å resolution and deposited in the PDB with accession code 5eh1. The electron-density map was interpreted for amino-acid residues 28–240, apart from two short loops. The IFN γ R2 structure revealed the fold common to other cytokine receptors: two fibronectin type III domains connected by a short linker. IFN γ R2 is a glycoprotein with five of the six potential N-linked glycosylation sites glycosylated, as confirmed by mass spectrometry and the crystal structure. Our analysis of glycosylation also uncovers the role of the oligosaccharide moieties at Asn110 and Asn137, which sandwich Trp131 and shield its hydrophobic aromatic ring from the solvent. Both potential disulfide bonds form but are not critical for the stability of IFN γ R2, as it is also stable in a reducing environment. The fifth cysteine Cys174 is bound to the monomeric cysteine.

Structure and sequence alignments revealed some important features of the 12 class 2 receptors. Their N-terminal D1 domains are more mutually similar than their C-terminal D2 domains (Fig. 3). D1 carries a distinctive so far unrecognized structural feature: a π -cation motif of sequentially distant stacked residues (X)WRWR(X) (KWRWRH in IFN γ R2;

Fig. 2a). Analysis of the receptor structures revealed three structurally highly variable regions (Fig. 4a), which most likely bring about binding specificity for their interacting partners. This hypothesis is further supported by the alignment of IFN γ R2 sequences from various species, which identified the highest sequence variability at positions coinciding with the structurally variable regions (Fig. 4b). An important structural feature distinguishing IFN γ R2 from the related IL10R2, gp130 and γ_c receptors is the specific positioning of the aromatic recognition epitope in IFN γ R2 (Fig. 2c).

We believe that the determination of the structure of the so far missing component of the interferon- γ signalling complex will enable a deeper understanding of the functioning of this important immunity cascade.

5. Related literature

The following references are cited in the Supporting Information for this article: Cancino-Díaz *et al.* (2002), Černý *et al.* (2007), Furche *et al.* (2014), O'Boyle *et al.* (2011) and Rüger *et al.* (2015).

Acknowledgements

This study was supported by the Czech Science Foundation (GA CR) grant No. 16-20507S and by project BIOCEV



OPEN ACCESS

 EDITED BY
 Laurent Roberto Chiarelli,
 University of Pavia, Italy

 REVIEWED BY
 Jeffery M. Tharp,
 Indiana University Bloomington,
 United States
 Richard Cooley,
 Oregon State University, United States
 Julia Shifman,
 Hebrew University of Jerusalem, Israel

 *CORRESPONDENCE
 Gustavo Fuertes,
 ✉ gustavo.fuertes@ibt.cas.cz,
 ✉ gustavo.fuertesvives@gmail.com

 RECEIVED 29 April 2023
 ACCEPTED 27 June 2023
 PUBLISHED 07 July 2023

 CITATION
 Pham PN, Zahradník J, Kolářová L,
 Schneider B and Fuertes G (2023),
 Regulation of IL-24/IL-20R2 complex
 formation using photocaged tyrosines
 and UV light.
Front. Mol. Biosci. 10:1214235.
 doi: 10.3389/fmols.2023.1214235

 COPYRIGHT
 © 2023 Pham, Zahradník, Kolářová,
 Schneider and Fuertes. This is an open-
 access article distributed under the terms
 of the [Creative Commons Attribution
 License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
 reproduction in other forums is
 permitted, provided the original author(s)
 and the copyright owner(s) are credited
 and that the original publication in this
 journal is cited, in accordance with
 accepted academic practice. No use,
 distribution or reproduction is permitted
 which does not comply with these terms.

Regulation of IL-24/ IL-20R2 complex formation using photocaged tyrosines and UV light

 Phuong Ngoc Pham^{1,2}, Jiří Zahradník^{3,4}, Lucie Kolářová¹,
 Bohdan Schneider¹ and Gustavo Fuertes^{1*}
¹Laboratory of Biomolecular Recognition, Institute of Biotechnology of the Czech Academy of Sciences, Vestec, Czechia, ²Faculty of Science, Charles University, Prague, Czechia, ³First Faculty of Medicine, BIOCEV Center, Charles University, Prague, Czechia, ⁴Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel

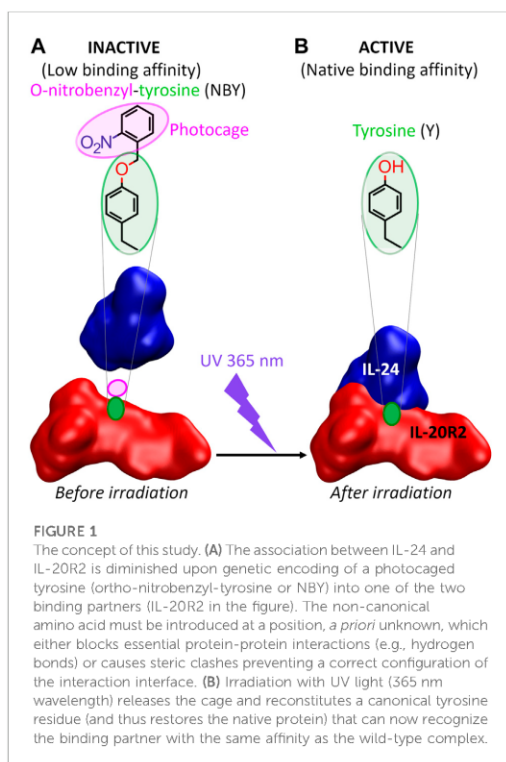
Human interleukin 24 (IL-24) is a multifunctional cytokine that represents an important target for autoimmune diseases and cancer. Since the biological functions of IL-24 depend on interactions with membrane receptors, on-demand regulation of the affinity between IL-24 and its cognate partners offers exciting possibilities in basic research and may have applications in therapy. As a proof-of-concept, we developed a strategy based on recombinant soluble protein variants and genetic code expansion technology to photocontrol the binding between IL-24 and one of its receptors, IL-20R2. Screening of non-canonical *ortho*-nitrobenzyl-tyrosine (NBY) residues introduced at several positions in both partners was done by a combination of biophysical and cell signaling assays. We identified one position for installing NBY, tyrosine70 of IL-20R2, which results in clear impairment of heterocomplex assembly in the dark. Irradiation with 365-nm light leads to decaging and reconstitutes the native tyrosine of the receptor that can then associate with IL-24. Photocaged IL-20R2 may be useful for the spatiotemporal control of the JAK/STAT phosphorylation cascade.

KEYWORDS

 protein-protein interactions (PPI), interleukin-24, cytokines, optobinders, genetically encoded non-canonical amino acids (ncAA), photocaged proteins, *ortho*-nitrobenzyltyrosine (NBY), photoxenoprotein engineering

1 Introduction

IL-24 is a multifunctional cytokine playing key roles in immune response, host defense, tissue homeostasis, and cell proliferation (Ma et al., 2011; Rutz et al., 2014; Liu et al., 2023). It is a member of a broader family of IL-10 related cytokines including IL-10, IL-19, IL-20, IL-22, IL-24, IL-26, IL-28 and IL-29 (Akdis et al., 2011). An increase in the expression levels of IL-24 is connected to autoimmune diseases, such as psoriasis (Kragstrup et al., 2008), inflammatory bowel disease (Andoh et al., 2009), and rheumatoid arthritis (de Melo et al., 2012). Moreover, a large number of studies suggest anticancer properties for IL-24, such as stimulation of apoptosis and autophagy, or inhibition of angiogenesis, invasion and metastasis (Menezes et al., 2018). Given the anti-oncogenesis effects of IL-24, it became a pharmacological target and even reached phase I clinical trials (Cunningham et al., 2005). Accumulated evidence strengthened the concept of the “bystander effect” according to which secreted IL-24, either in normal or cancer cells, induces tumor apoptosis in presence of IL-20/IL-22 receptors (Su et al., 2005).



Membrane-bound receptors for IL-24 comprise IL-22R1, IL-20R1, and IL-20R2. IL-24/receptor interactions are mediated by the extracellular domains of these receptors. IL-24 signals through both the IL-20 receptor (IL-20R1/IL-20R2) and IL-20R1/IL-20R2 heterodimers (Dumoutier et al., 2001; Parrish-Novak et al., 2002; Wang et al., 2002). Upon binding and heterotrimer formation, IL-24 triggers a signaling cascade in the target cells through Janus Kinase/Signal Transducer and Activator of Transcription (JAK/STAT) pathway, which subsequently activates downstream transcription factors, like STAT3 and STAT1, through phosphorylation (Wang and Liang, 2005).

Since many physiological and pathophysiological roles of IL-24 critically depend on its interaction with cognate receptors (Wang et al., 2002; Wang et al., 2004), we hypothesize that controlling the binding between IL-24 and the shared receptor IL-20R2 could be useful for both basic research and therapeutics. The use of light to switch ON and OFF protein-protein interactions offers superior capabilities in terms of temporal and spatial resolution (Hoorens and Szymanski, 2018). We speculate that assays based on photocontrolled interleukins and/or their cognate receptors may be applied to the detection of specific interleukins, the screening of anti-interleukin/anti-receptor antibodies, or the screening of JAK/STAT inhibitors.

From the molecular engineering point of view, there are two major approaches to control protein functions by light

(Kneutinger, 2022). The first is hybrid protein optogenetics, where a chimera between an intrinsically light-responsive protein and a target protein is made. The second is photoxenoprotein engineering, where the target protein is modified at defined residues with light-responsive non-canonical amino acids (ncAA). NcAA can be incorporated into proteins by genetic code expansion technology through an orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA_{CUA} pair (Manandhar et al., 2021). Several photocontrolled ncAA have been genetically encoded in *Escherichia coli* and mammalian cells, including analogs of tyrosine, lysine, cysteine, serine and histidine (Wu et al., 2004; Deiters et al., 2006; Lemke et al., 2007; Gautier et al., 2010; Arbely et al., 2012; Nguyen et al., 2014; Hauf et al., 2017; Luo et al., 2017; Baumann et al., 2019; Cheung et al., 2023). We choose to photocage tyrosines in the form of *ortho*-nitrobenzyl-tyrosine (NBY) for two reasons. First, NBY has been proposed as a universal proximal cage for the temporal blockage of protein activity (Wang et al., 2019; Wang et al., 2021). Second, NBY has been successfully employed to photocontrol antigen-antibody interactions (Bridge et al., 2019; Jedlitzke et al., 2019; Joest et al., 2021; Jedlitzke and Mootz, 2022; O'Shea et al., 2022; Yilmaz et al., 2022; Bridge et al., 2023).

Thus, as a first step towards the photocontrol *in vivo* of cell signaling pathways that rely on IL-24-mediated complex formation, we designed and successfully tested the following *in vitro* approach with purified proteins recombinantly expressed in *E. coli* (Zahradnik et al., 2019) (Figure 1). Incorporation of photocaged NBY residues at positions critical for molecular recognition into either IL-24 or IL-20R2 would keep the protein partners incompetent for binding and unable to convey the signal. Following the removal of the cage by UV irradiation, the native protein structure and activity (binding and signaling) would be restored resulting in complex formation. In the future, the ability to switch OFF the binding between IL-24 and IL-20R2 in the dark and switch it ON by UV light may find applications in immunology and possibly also in therapeutics.

2 Materials and methods

2.1 Miscellaneous chemicals

pEVOL-NBY was a gift from Dr. Petr Schultz (Deiters et al., 2006). NBY was purchased from Accela. For a more complete list of reagents, please refer to Supplementary Table S0.

2.2 Gene cloning

An engineered IL-24 variant named IL-24B4 had been sequence-optimized to achieve soluble expression in *E. coli* following structural bioinformatics design (Zahradnik et al., 2019). This sequence contains 29 mutations relative to wild-type IL-24 (Supplementary Table S1, residue numbering according to UniprotKB Q13007). A pHisSumo vector containing such interleukin-24 variant with an N-terminal HisSumo tag (Zahradnik et al., 2019) was used as the starting material for cloning. We used the primers 1 and 2 (Supplementary Table S2) to sandwich the IL-24 gene between an N-terminal HisSumo-tag and a

C-terminal Strep-tag. The final construct was labeled as pHisSumo_IL-24B4_CStrep (Supplementary Figure S1A).

An IL-20R2 variant sequence, referred to as IL-20R2D, which was also optimized to be expressed and soluble in *E. coli*, contained 23 mutations relative to the wild-type (Supplementary Table S1 residue numbering according to UniprotKB Q6UXL0). The design details can be found in the main text and Supplementary Note S1. The ordered DNA string consists of a HisSumo tag and a Strep tag flanking IL-20R2 at the N-terminus and C-terminus, respectively. The final construct was labeled as pHisSumo_IL-20R2D_CStrep (Supplementary Figure S1B).

TAG triplets to incorporate non-canonical amino acids were introduced in the plasmid sequences by standard site-directed mutagenesis using primers 5 to 11 (Supplementary Table S2).

2.3 Protein expression

To express proteins with canonical amino acids, the pHisSumo_IL-24B4_CStrep and pHisSumo_IL-20R2D_CStrep plasmids were transformed into *E. coli* BL21 AI strain. The bacterial cells were grown in TB medium containing 1.2% tryptone, 2.4% yeast extract, 0.5% glycerol, 33.7 mM Na₂HPO₄, 22.0 mM KH₂PO₄, 8.55 mM NaCl, 18.7 mM NH₄Cl, 2 mM MgSO₄/MgCl₂, and metal mix. The antibiotic kanamycin was supplied to the medium and cells were grown at 30°C with shaking at 200 rpm. When bacterial OD₆₀₀ reached 0.6 to 0.8, cells were induced with IPTG and arabinose at 0.1 mM and 0.02% v/v final concentration, respectively, and grown at 16°C for another 20 h.

For proteins containing non-canonical amino acids, plasmids encoding the proteins-of-interest were co-transformed in BL21 AI with pEVOL-NBY, a plasmid encoding a tyrosyl tRNA synthetase derived from *Methanocaldococcus jannaschii* specific for *ortho*-nitrobenzyl-tyrosine (MjNBYS) and its cognate suppressor tRNA (tRNA_{CUA}) (Deiters et al., 2006).

Bacteria were grown at 30°C and shaking at 200 rpm in TB medium described above including antibiotics kanamycin and chloramphenicol. During culturing, we prepared a 50 mM stock of NBY by resuspending the amino acid in 50% DMSO heated at 70°C and adding NaOH dropwise until a homogeneous solution was obtained. When bacterial OD₆₀₀ reached between 0.4 and 0.5, NBY was added to the culture at a final concentration of 1 mM. After 25 more minutes, the temperature was decreased to 16°C, and cells were induced by both IPTG and arabinose at 0.1 mM and 0.02% final concentration, respectively. Overnight-grown bacteria (20 h at 16°C) were separated from the medium culture by centrifugation at 5,000 g for 10 min at 4°C. All cell pellets were stored at -80°C until needed.

2.4 Protein purification

The cell pellet was suspended in cold washing buffer 50 mM Tris, pH = 8, 100 mM NaCl and sonicated 2 s on, 2 s off during 5 min with 50 W power on ice. The fraction of soluble proteins was collected by centrifuging at 40 000 g for 20 min at 4°C and the supernatant was passed over streptactin XT agarose beads equilibrated at room temperature by 50 mM Tris, pH = 8,

150 mM NaCl, 1 mM EDTA. After brief incubation for the proteins containing Strep tag to bind to the beads, the unbound proteins were washed with the same buffer. Commercial elution buffer BXT (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM biotin) was used to elute the bound Strep-tagged proteins. All the eluted proteins were further purified by size-exclusion chromatography (SEC) using a Superdex 75 Increase 10/300 GL column (GE Healthcare, Chicago, IL, United States) equilibrated with buffer 50 mM Tris, pH = 8, 100 mM NaCl.

IL-24B4 was digested with Sumo protease to remove the N-terminal HisSumo tag by at a ratio 1:200 (protease:protein) at room temperature overnight (16 h). The digestion mixture was passed over nickel-NTA agarose beads. Uncleaved IL-24 and the cleaved HisSumo fragment were retained on the beads, while cleaved IL-24 without the N-terminal HisSumo tag was recovered from the flow-through. Cleaved IL-24 was subsequently purified by SEC as described before, pooled and concentrated.

2.5 Protein characterization

The quality and quantity of all protein samples was checked by electrophoresis, UV/Visible spectroscopy, and mass spectrometry (de Marco et al., 2021). Proteins were analyzed by 15% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Blue. UV/Vis spectra were taken in a Nanodrop spectrophotometer (DeNovix) using 10 mm pathlength cuvettes. The protein extinction coefficient at 280 nm, estimated from the sequence using web server (<https://web.expasy.org/protparam/>) was used to calculate protein concentration based on Beer-Lambert law.

For mass spectrometry analyses, proteins were diluted with 100 µL of 5% acetic acid in water and loaded onto Opti-trap C4 cartridge (Optimize Technologies), washed 4 times with 250 µL of 5% acetic acid in water and eluted with 100 µL of 80% acetonitrile, 5% acetic acid. Proteins were analyzed by direct infusion using syringe pump at a flow rate 3 µL/min connected with an electrospray ion source of TimsTOF Pro mass spectrometer (Bruker Daltonics). The mass spectrometer was externally calibrated using 0.1% (w/w) sodium trifluoroacetate. Proteins were measured in positive mode. The data were processed using SNAP algorithm—a part of DataAnalysis 5.3 software (Bruker Daltonics). Deconvoluted spectra were generated using the UniDec 5.0.2 software (University of Oxford) (Marty et al., 2015). IL-20R2D and IL-20R2D Y70NBY were also measured on a high resolving power and high accuracy 15T SolariX XR FT-ICR mass spectrometer (Bruker Daltonics). Expected protein masses (average and monoisotopic) were calculated with 2M data acquisition, by submitting the protein sequences to the website https://web.expasy.org/compute_pi/.

The circular dichroism spectra were measured using a 0.1 cm quartz cuvette on ChirascanTM-plus spectrometer (Applied Photophysics). Proteins were diluted in buffer (Tris 50 mM NaCl 100 mM pH = 8.0) to final concentration of ~0.15 mg/mL. The data were recorded in the range of wavelengths of 190–260 nm with a step of 1 nm and time-per-point of 1 s at room temperature. The resulting spectra were water-subtracted and normalized to the concentration of the corresponding sample and are given as molar circular dichroism ($\Delta\epsilon$) vs wavelength. The data were

processed by CDNN 2.1 software and BestSel (<https://bestsel.elte.hu/index.php>) (Micsonai et al., 2022).

Label-free protein unfolding assays were done by detecting the temperature-dependent change in tryptophan fluorescence at emission wavelengths of 330 and 350 nm (Alexander et al., 2014). Melting temperatures were determined from the maximum of the first derivative of the fluorescence ratios (F350/F330). For thermal unfolding experiments, the proteins were diluted to a final concentration of ~0.15 mg/mL. The samples were loaded into UV capillaries (NanoTemper Technologies) and experiments were carried out using the Prometheus NT.48.

2.6 UV-decaying

200 μ L of the proteins at 200 nM concentration were transferred to 1 mm path-length cylindrical cells (Hellma) and irradiated at 365 nm with an LED (Thorlabs M365L2) attached to a collimator (Thorlabs SM2F32-A) at 100 mW of power for 5 min at room temperature. Irradiated proteins were used immediately.

2.7 Affinity measurements by MST

We determine the binding affinity (dissociation constant) by MicroScale Thermophoresis (MST) using a Monolith NT.115 instrument (NanoTemper Technologies) (Jerabek-Willemsen et al., 2014). 200 nM of IL-20R2D, IL-20R2D Y70NBY, IL-20R2D Y74NBY, or IL-20R2D Y70NBY/Y74NBY were labeled with an NTA-conjugated fluorescent dye as recommended by the kit His-Tag Labeling RED-tris-NTA second generation. Binding assays were performed in buffer 50 mM Tris, pH = 8, 100 mM NaCl, 0.02% pluronic F127. The instrument settings were 20% MST power and 40% LED power. Data from triplicate experiments were fitted to mass-action kinetics by using the NT Analysis software version 1.5.41:

$$F(c) = A + B \left(\frac{c + c_T + K_d - \sqrt{(c + c_T + K_d)^2 - 4c c_T}}{2c_T} \right) \quad (1)$$

Where $F(c)$ is the fluorescence signal as a function of concentration; c is the concentration of IL-24B4 (unlabeled) c_T is the concentration of IL-20R2D variants (fluorescently labeled), which was held constant at 50 nM; and K_d is the dissociation constant; A is the titration curve baseline; B is the titration response range. To facilitate the comparison between datasets with distinct values of A and B , the fraction of bound molecules (normalized between zero and one) was calculated as:

$$\text{Fraction bound} = \frac{F(c) - A}{B} \quad (2)$$

2.8 Affinity measurements by yeast display

Yeast surface display of IL-24B4 and subsequent binding experiments were performed by using the enhanced yeast display platform pFYDC1 plasmid (Addgene ID: 162458) and

Saccharomyces cerevisiae EBY100 strain (Zahradnik et al., 2021a). The bait, an IL-24B4 version including an N-terminal linker of 17 amino acids, was cloned between *Nde*I and *Bam*HI sites (see full construct sequence in Supplementary Table S1), and transformed into yeast by lithium acetate method. Protein expression was achieved by cultivation in galactose-rich expression media 1/9 at 20°C for 24 h (Zahradnik et al., 2021a). Prey proteins (IL-20R2D, IL-20R2D Y70NBY and their UV irradiated versions) were labeled by using CF[®]640R succinimidyl ester dye (Biotium) in 1:3 M excess.

The binding affinities of labeled IL-20R2D versions to IL-24B4 expressed on yeast surface were determined by flow cytometry titration experiments. We used 5 different concentrations (500, 100, 20, 4, and 0.1 nM) of labeled prey protein in PBS buffer supplemented by 1 g/L of Bovine Serum Albumin Fraction V (PBSB buffer, Sigma-Aldrich, Cat#10735078001), and 10 nM bilirubin (Sigma-Aldrich, Cat#14370) for activation of eUnaG2 fluorescence (excitation at 498 nm, emission at 527 nm). To increase the control over the measurement, we enriched the suspension by addition of 0.5% of IL-24S, an engineered clone featuring 300 times tighter binding, which was obtained by yeast display affinity maturation (Supplementary Table S1 and Supplementary Figure S2).

Binding suspensions were incubated overnight at 4°C, washed twice with ice-cold PBSB and analyzed by using CytoFLEX S Flow Cytometer (Beckman Coulter, United States, Cat# N0-V4-B2-Y4). The gating and analysis strategies were described previously (Zahradnik et al., 2021a; Zahradnik et al., 2021b). Briefly, the unspecific binding signals were subtracted from the binding signals and a Eq. 1 was fitted to the data by nonlinear least-squares regression using Python v3.7 (<https://www.python.org>).

2.9 JAK/STAT signaling assays in human cells

HeLa cells were grown in a media containing Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1% of PenStrep and 1% of sodium pyruvate. To ensure sufficient amounts of the receptors IL-20R2 and IL-22R1, two plasmids encoding the native sequences (Supplementary Table S1) were constructed (oligos 12 to 15 of Supplementary Table S2) and co-transfected in HeLa cells (0.66 μ g of IL-20R2 and 1.33 μ g of IL-22R1) using the jetPrime transfection reagent (Polyplus 114–07) following manufacturer's instructions. The next day, the cells were treated for 30 min with the obtained interleukin variants at the indicated concentrations. After the incubation period, the plates were put on ice, the media discarded, and cells were washed with cold PBS buffer. Then, 100 μ L of RIPA lysis buffer (Tris 50 mM pH = 8.0, NaCl 150 mM, EDTA 5 mM, IGEPAL CA-630 1%, sodium deoxycolate 10%, SDS 10%) supplemented with protease and phosphatase inhibitors was added to each well. Cells were scraped and put into Eppendorf tubes for 20 min on ice, followed by centrifugation at 4°C, 18,000 g for 30 min. Protein concentration in the supernatant fraction was quantified by the BCA assay (QPRO BCA kit from Cyanagen) using BSA as the standard. Equal protein amounts were loaded on SDS-PAGE gels. Western blotting to nitrocellulose membranes was done in a Transblot Turbo Bio-rad system. For immunoblotting we use two primary antibodies:

anti-phospho-STAT3 (Tyr705, D3A7) monoclonal antibody raised in rabbit, and anti α -tubulin monoclonal antibody raised in mouse (Sigma). As secondary antibodies we used anti-mouse or anti-rabbit peroxidase-conjugated IgG. One milliliter of peroxidase substrate was added to the membranes and the resulting chemiluminescence was imaged using an Azure400 system. All original Western blot images can be found in the Supplementary Zip file. Band quantification was done in ImageJ (<https://imagej.nih.gov/ij/index.html>) as previously described (Stael et al., 2022). The band intensities of pSTAT3 were divided by the corresponding band intensities of tubulin and subtracted from a control experiment (untreated cells). Normalized and background-subtracted band intensities were plotted as a function of the concentration of added variant and a dose-response curve was fitted to the data:

$$\frac{pSTAT3}{tubulin} = \frac{Y_{max}}{1 + (EC_{50}/c)^{n_H}} \quad (3)$$

Where Y_{max} is the maximum observed signal, EC_{50} is the half maximal effective concentration, c is the concentration of added interleukin (IL-24B4, IL-20R2D and their NBY counterparts), and n_H is the Hill coefficient (which was fixed to 1, i.e., non-cooperative binding), to prevent overfitting due to the reduced number of points).

3 Results

3.1 Protein design combining canonical and non-canonical amino acid mutagenesis

In vitro testing of photo-induced changes in binding affinity requires pure and active protein partners. We designed engineered variants that are more amenable to downstream manipulations and may be used as surrogates of wild-type proteins after a scrutiny of their performance.

3.1.1 Enhancing solubility and stability with canonical amino acids

In our previous work, we successfully optimized the sequence of IL-24 based on canonical amino acid mutagenesis to achieve soluble expression in *E. coli* hosts (Zahradník et al., 2019). Here we have applied a similar approach to the high affinity receptor of IL-24, the IL-20R2. In order to obtain an stabilized soluble extracellular portion of IL-20R2 protein, we employed PROSS algorithm (Goldenzweig et al., 2016) (Supplementary Note S1, Supplementary Table S1, Supplementary Figure S2, and Supplementary Figure S3A). We then tested the expression of our newly designed IL-20R2D in *E. coli* BL21 (DE3) under the control of T7 promoter and compared it to IL-20R2 wild-type under the same conditions. Unlike the wild-type gene, our construct is expressed as a soluble protein in large quantities ~ 20 mg.L⁻¹ (Supplementary Figure S3B). Subsequent CD and melting measurements showed evidence of a folded protein (Supplementary Figure S3C) with a melting temperature of $\sim 45^\circ\text{C}$ (Supplementary Figure S3D). The protein showed considerable stability and retention of binding to IL-24B4 (Zahradník et al., 2019).

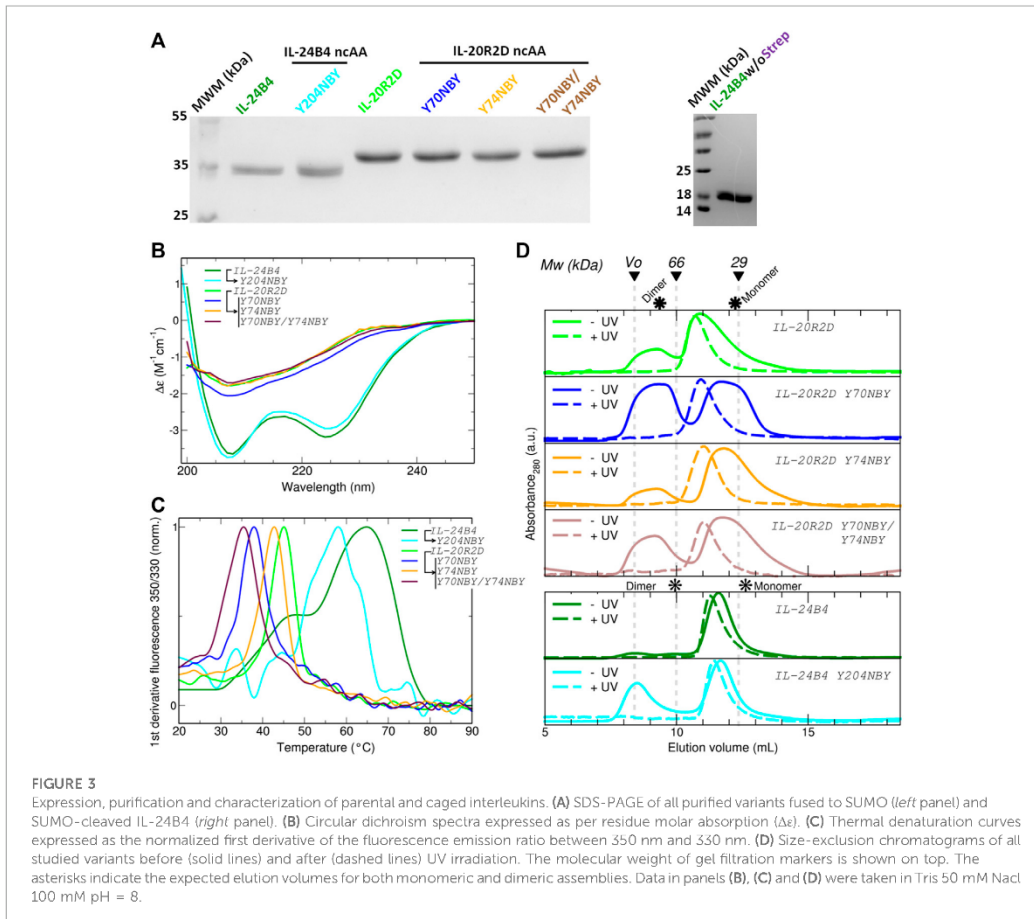
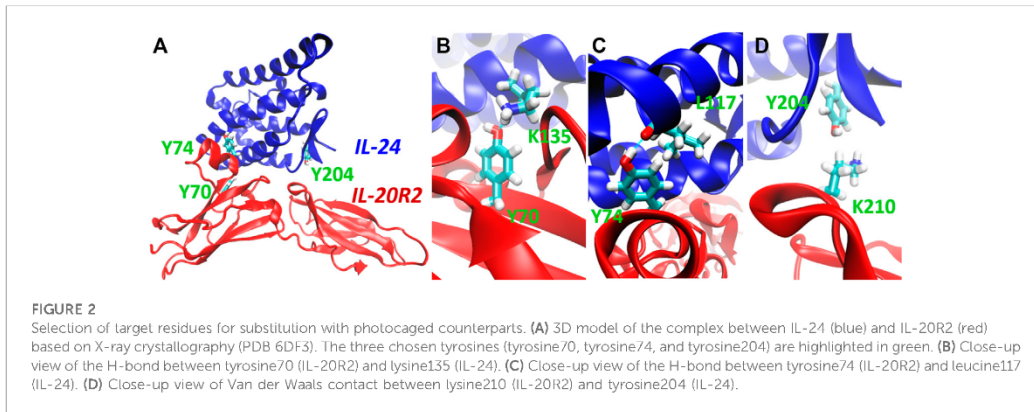
3.1.2 Adding photocontrol with photocaged ncAA

As mentioned before, we chose to photocage tyrosine residues with the *o*-nitrobenzyl moiety. We then looked for tyrosine residues at the binding interface between IL-24B4 and IL-20R2D. Since no high-resolution structure exists of the complex between the variants used in the present study, we used the X-ray crystal structure of the homologous native ternary complex (IL-24/IL-22R1/IL-20R2, PDB ID: 6DF3) (Lubkowski et al., 2018) (Figure 2A). The structure suggests that the main contribution to complex formation are the contacts between IL-24 and IL-20R2. The IL-24/IL-20R2 interface is shown schematically in 2D (Supplementary Figure S4A). The interface area covers ~ 900 Å² and is stabilized by several forces including 13 hydrogen bonds (H-bonds), 2 potential salt bridges, and a number of polar, water-mediated and hydrophobic contacts. We identified three tyrosine candidates: one in IL-24 (Y204), and two in IL-20R2 (Y70 and Y74). The phenolic hydroxyl group of tyr70 in IL-20R2 makes a hydrogen bond (H-bond) with the ϵ -amino group of lys135 in IL-24 (Figure 2B). The sidechain OH of tyr74 in IL-20R2 is H-bonded to the backbone CO group of leu117 in IL-24 (Figure 2C). Finally, tyr204 in IL-24 makes Van der Waals interactions with lys210 of IL-20R2 (Figure 2D) and is also involved in H-bond interactions with IL-22R1 (Supplementary Figures S4B–E). Considering all these data, we prepared three single mutants: IL-24B4 Y204NBY, IL-20R2D Y70NBY and IL-20R2D Y74NBY. In addition, we prepared the double mutant of IL-20R2D Y70NBY/Y74NBY.

3.2 Production and characterization of photocaged interleukins and receptors

For the design of the final constructs, we considered one additional phenomenon. Codon decoding by orthogonal aminoacyl-tRNA synthetase/trNA_{CUA} pairs is not 100% efficient, i.e., protein translation can also terminate at the introduced codon resulting in a mixture of truncated and full-length proteins that can be difficult to separate. To avoid such a problem, we added a C-terminal Strep-tag. All our final constructs express target proteins containing a HisSumo tag at the N-terminus and a Strep-tag at the C-terminus. First, we checked whether addition of the Strep tag alters the yield of IL-24B4 and IL-20R2D. The purity was assessed by SDS-PAGE (Figure 3A). In the case of IL-24B4, the HisSumo tag was removed to avoid interference with downstream affinity measurements by microscale thermophoresis (MST). This is because for MST, the His-tagged target protein is non-covalently labeled with a NTA-conjugated fluorescent dye. Mixing with another protein containing a His-tag will cause the dye to re-equilibrate between the two binding partners thus confounding the experiment. Sumo-free IL-24B4 is shown in Figure 3A (right panel).

To obtain proteins with NBY incorporated at specific positions, an evolved orthogonal aminoacyl tRNA-synthetase (MjNBYRS) and suppressor tRNA (trNA_{CUA}) pair was added to the native bacterial translation system (Deiters et al., 2006). Moreover, the bacteria cells were fed with the dissolved non-canonical amino acid NBY in TB culture medium. The expressed full-length non-canonical receptors were purified by one-step Strep-tag procedure as done for the parental proteins. In addition, the negative controls without NBY



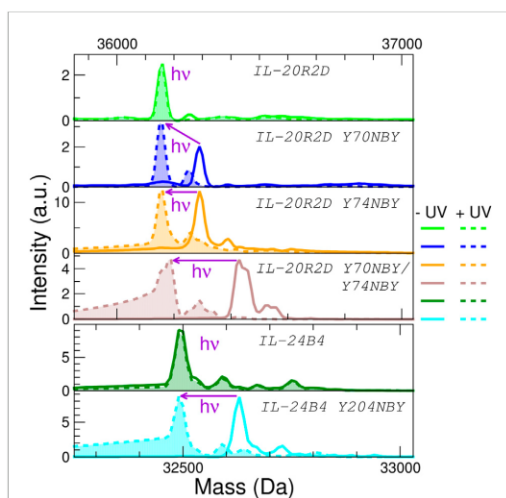


FIGURE 4

Analysis of the effect of protein UV irradiation by mass spectrometry. Deconvoluted mass spectra of the six interleukin variants used in this study before (solid lines) and after (dashed lines with filled areas) exposure to light ($\lambda = 365$ nm, 5 min at 100 mW). No significant mass changes are observed for the parental proteins (IL-20R2D and IL-24B4) upon UV irradiation. NBY-containing IL-20R2D and IL-24B4 are converted back to IL-20R2D and IL-24B4, respectively, by UV light. Average masses can be found in Supplementary Table S3.

supplement were performed for size comparison. Clear bands at the expected molecular weight were found in the case IL-24B4 Y204NBY (Supplementary Figure S5A) IL-20R2D Y70NBY, IL-20R2D Y74NBY (Supplementary Figure S5B) in the presence of ncAA. On the contrary, for the samples without added NBY there were essentially no proteins (Supplementary Figure S5). These results suggest that the ncAA NBY was incorporated into the interleukins and their receptors at selected residues only in *E. coli* cells containing both the non-canonical amino acid and the orthogonal translation machinery.

To check whether NBY was indeed present in the purified polypeptides, mass spectrometry was employed (Figure 4). The experimentally determined masses are very close to the theoretical masses (plus/minus 2 Da) (Supplementary Table S3). We did not detect mis-incorporation of phenylalanine residues or reduction of the nitro group, two common issues associated to the genetic encoding of NBY (Koch et al., 2022). These results confirm unambiguously that the non-canonical amino acid NBY was introduced into the target proteins at high levels.

We further characterized the obtained protein variants by circular dichroism (CD) and differential scanning fluorimetry (DSF). The CD spectra of the IL-24B4 variants showed two minima at ~ 208 nm and ~ 224 nm (Figure 3B). In opposition, the CD spectra of the IL-20R2D variants featured a single minimum around 207 nm (Figure 3B). Analysis of secondary structure content revealed that IL-24B4 variants were enriched in α -helices while IL-20R2D mutants contained predominantly antiparallel β -sheets

(Supplementary Table S4) in agreement with their crystal structures. Analysis of melting temperatures by DSF indicated significant differences in thermal stability among the mutants (Figure 3C). IL-24 variants were more stable than IL-20R2D variants, $\sim 60^\circ\text{C}$ vs. $\sim 40^\circ\text{C}$, respectively (Supplementary Table S5). In both cases, NBY reduced the protein's melting temperature. The largest effect was observed in the double mutant (Y70NBY/Y74NBY) that had $\sim 10^\circ\text{C}$ lower melting temperature than the parental protein. Therefore, although the introduced NBY caused minimal structural changes, protein stability was reduced.

3.3 Testing the effect of UV irradiation

Supplementary Table S6 shows the employed UV wavelengths together with protein's UV/Visible absorption spectra. Prior to measuring protein affinities and cellular signaling, we checked the effect of UV light on protein oligomerization and molecular weight by size-exclusion chromatography (SEC) and mass spectrometry. According to SEC (see chromatograms in Figure 3D), the apparent molecular weights of all studied variants lay in between the expectation for a monomer and a dimer (Supplementary Table S6). Therefore, at least two interpretations are possible. Either the proteins are in monomer/dimer equilibrium or they feature an expanded ("open") monomeric conformation. The NBY mutations tend to make the proteins more aggregation prone and substantial protein amounts were found in the void volume. Importantly, no major changes in the elution profiles were observed before/after UV and no large aggregates were found upon UV exposure. However, UV-irradiated samples eluted earlier than the non-irradiated counterparts (Supplementary Table S6). Although partial UV-induced dimerization cannot be completely ignored, we conclude that a molecular interaction with a 1:1 stoichiometry is a suitable analytical model.

Second, we ruled out any potential UV-induced photodamage to the proteins. The masses of IL-24B4 and IL-20R2D before and after illumination were virtually identical, suggesting that the irradiances used in this study are insufficient to cause significant residue modifications (Figure 4). Next we checked the efficiency of decaging. We shined UV light on the 4 NBY-containing variants and observed that the mass decreased ~ 135 Da (or ~ 270 Da in the case of the double mutant), which corresponds to that of the nitrobenzyl moiety, making it indistinguishable from the mass of canonical variants (Figure 4). For better visualization of the mass difference we also show detailed mass spectra of 35+ charge state (Supplementary Figure S7).

Thus, we prove that all photocaged interleukins/receptors can be efficiently decaged by UV irradiation, thereby regenerating the parental proteins natively containing a tyrosine residue (or residues in the case of the double mutant).

3.4 Measuring interleukin-receptor interactions and their effects

Having found conditions for efficient decaging, we next monitored the effect of the produced interleukin variants on IL-24/IL-20R2 complex formation in three environments of increasing

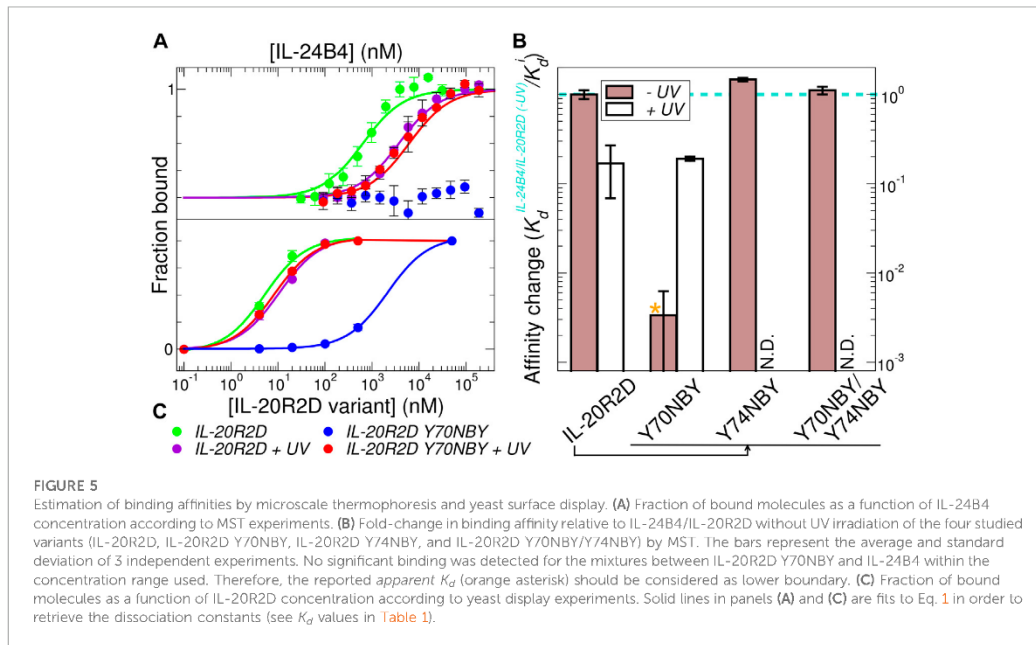
TABLE 1 Binding affinities between IL-24B4 and IL-20R2D variants. Dissociation constant (K_d) and half-maximal effective concentration (EC_{50}) values are expressed as the mean \pm s.d. (3 biological replicates for MST, 3 technical replicates for yeast display, and 2 technical replicates for Western blot).

Partner 1 ^a	Partner 2	Method	MST	Yeast display	Cell signaling (pSTAT3)
			K_d (nM)	K_d^b (nM)	EC_{50}^c (nM)
IL-24B4	IL-20R2D ^c		630 \pm 140	5.0 \pm 0.6	0.28 \pm 0.07
IL-24B4 Y204NBY	IL-20R2D ^c		N.P.	N.D.	0.9 \pm 0.5
IL-24B4 Y204NBY + UV	IL-20R2D ^c		N.P.	N.D.	0.76 \pm 0.24
IL-24B4	IL-20R2D + UV		3,700 \pm 400	9.6 \pm 1.6	N.P.
IL-24B4	IL-20R2D Y70NBY		>94,000	~2000	N.P.
IL-24B4	IL-20R2D Y70NBY + UV		3,300 \pm 600	7.6 \pm 0.4	N.P.
IL-24B4	IL-20R2D Y74NBY		430 \pm 30	N.D.	N.P.
IL-24B4	IL-20R2D Y70NBY/Y74NBY		570 \pm 60	N.D.	N.P.

^aFor microscale thermophoresis (MST) experiments the partner 2 was held constant and partner 1 titrated, while the opposite is true for yeast display experiments.

^b K_d values obtained from yeast display should be considered *apparent* because they reflect not only affinity but also avidity due to the high protein densities on the cell surface.

^cProduction of phosphorylated STAT3 (pSTAT3) as a function of added partner 1. Note that in this case the partner 2 corresponds to IL-20R2/IL-22R1, expressed in HeLa cells. N.D., means not determined; N.P., indicates a measurements that it is not possible due to low sample amounts or limitations of the cell assays (see main text more details).



complexity (Supplementary Figure S8). First, with both interacting partners free in solution by MST. Second, with one binding partner immobilized on the yeast cell surface and the other in solution by yeast display. Finally, with the receptors natively inserted in the membrane of human cells and the others exogenously added to the growth media. The first two methods directly report on IL-24/IL-20R2 binding affinities, while the latter is sensitive to the impact that

such interactions have on signal transduction through the JAK/STAT pathway.

3.4.1 MST measurements

We determined binding affinities by MST, which measures temperature-induced changes in fluorescence intensity, using fluorescently labeled variants at a fixed concentration and mixing

with variable amounts of unlabeled binding partners. Unfortunately, the obtained amounts of IL-24B4 204NBY were too low. Hence, we concentrated on NBY-containing mutants of IL-20R2D for which we obtained sufficient yields for biophysical studies. The dissociation constant (K_d) of the complex between IL-24B4 and IL-20R2D is $\sim 0.5 \mu\text{M}$ (see K_d values in Table 1 and raw data in Supplementary Figure S9) in agreement with the literature (Zahradník et al., 2019). Exposure to UV light slightly diminished the binding affinity by a factor of 6 (compare green and violet traces in Figure 5A). Of the three tested variants containing photocaged non-canonical amino acids, IL-20R2D Y74NBY and the double-mutant IL-20R2D Y70NBY/Y74NBY bound to IL-24B4 with similar affinities as their canonical counterparts (Figure 5B). Non-irradiated IL-20R2D Y70NBY showed virtually no binding to IL-24B4 within the concentration range used (Figure 5A), suggesting a decrease of binding affinity by a factor of at least 300 relative to the canonical complex (Figure 5B). Gratifyingly, photoconversion of IL-20R2D Y70NBY back to IL-20R2D by UV light shifted the binding curve towards the values observed for the canonical protein (Figure 5A). Thus, we found one residue position in IL-20R2D, Y70, where NBY significantly blocks the heterodimerization with IL-24B4. In such a case, complex formation can be observed after cage photolysis by UV light. Unexpectedly, the double-mutant displayed high binding affinity in the absence of UV irradiation suggesting non-additive effects of the two mutations where the inhibitory effect of Y70NBY is somehow compensated for by Y74NBY.

3.4.2 Yeast display measurements

To support the previous results in a more cellular context, we also estimated binding affinities using yeast display. The four purified and fluorescently labeled IL-20R2D variants (canonical and non-canonical, with and without UV stimulation) were incubated at different concentrations with yeast cells expressing on their surface IL-24B4 fused to a fluorescent protein. The fluorescent signal arising from each binding partner was then analyzed by flow cytometry (Supplementary Figure S10). In these plots, binding events appear along the diagonal. UV illumination did not cause a significant change (2-fold reduction) in the population of IL-20R2D/IL-24B4 complex (Figure 5C). In the case of IL-20R2D Y70NBY in the absence of UV light, complex formation with IL-24B4 was largely attenuated within the concentration range used (Figure 5C). Upon illumination, the number of heterodimers increased and reached similar levels as for the canonical complex (Figure 5C). The apparent K_d values, including both affinity and avidities, are listed in Table 1.

When comparing the relative trends, both MST and yeast display agree in suggesting low levels of complex formation between IL-20R2D Y70NBY and IL-24B4 in the absence of light. However, the binding affinities estimated by yeast display were higher (approximately 100-fold) than those determined by MST in diluted protein solutions, probably due to avidity (multiple interaction events) and crowding effects. In summary, the two methods point to two completely different binding scenarios. On one hand, low binding affinity of IL-20R2D Y70NBY (caged receptor) against IL-24B4 in the dark. On the other hand, high binding affinity (similar to the parental proteins) upon UV illumination and subsequent decaging.

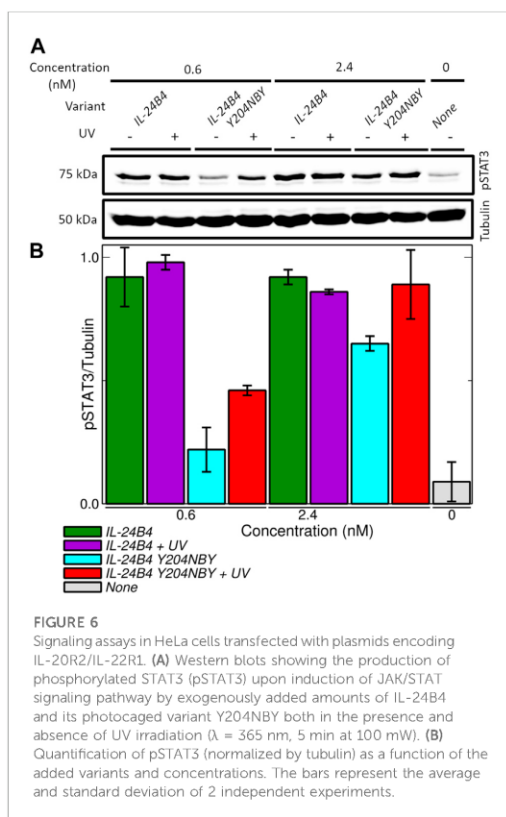
3.4.3 Cell signaling assays

Eventually, we measured the ability of our engineered variants IL-20R2D and IL-24B4 to activate the JAK/STAT signaling cascade in a human cell line (HeLa). To this end, we quantified the levels of the transcription factor STAT3 phosphorylated at tyrosine705 (pSTAT3 for short), which is a well-known marker implicated in the signaling of IL-24 and other cytokines (Darnell et al., 1994). pSTAT3 levels after 30 min exposure with canonical or non-canonical interleukin variants were determined by Western blotting using an anti-pSTAT3 antibody and subsequent enhanced chemiluminescence detection.

Initially, we checked whether our engineered versions (IL-24B4 and IL-20R2D) and the endogenous levels of native receptors (IL-20R2 and IL-22R1) present in HeLa cells can yield detectable amounts of pSTAT3. As a positive control we used interferon- α (IFN α), a cytokine that is known to induce rapid changes in JAK/STAT phosphorylation and initiate the signaling (Pellegrini et al., 1989). Little production of pSTAT3 was found in non-transfected cells regardless of the presence of IL-24B4 (Supplementary Figure S11A). On the contrary, significant amounts of pSTAT3 were found in transfected cells upon addition of IL-24B4 (Supplementary Figure S11A) suggesting that HeLa cells express limited amounts of endogenous receptors and require a boost of IL-24 receptors to enhance signal-to-noise. Similar amounts of pSTAT3 were found in the case of interferon- α stimulation (Supplementary Figure S11B). To test the capability of IL-20R2D to transduce biological signals, we created a chimeric receptor (IL-20R2DC) where the extracellular portion of full-length IL-20R2 was replaced by the designed IL-20R2D (Supplementary Table S1). The results (Supplementary Figure S11C) support the notion that IL-20R2DC can bind IL-24B4 and trigger signaling.

Surprisingly, we found that IL-20R2D signals through the same receptors and pathway although at high protein concentrations (Supplementary Figure S12A). Signal transduction through IL-20R2 was previously observed for IFN α 2 (Pattyn et al., 1999). Such a background level of pSTAT3 induction by IL-20R2D would complicate the interpretation of the results when both IL-24B4 and IL-20R2D are present. Taken together, these results confirm that the designed variants, IL-24B4 and IL-20R2D, recapitulate the role of native proteins, i.e., IL-24B4 binds its cognate membrane-bound receptors (IL-20R2, IL-22R1) and initiates the JAK/STAT pathway. However, a limitation of our cell signaling assays is that they do not unambiguously inform on the effect of IL-20R2D.

Therefore, we focused on the IL-24B4 Y204NBY mutant, which could be purified in sufficient amounts for the cell assays. We measured the intracellular levels of pSTAT3 reporter after incubation with different concentrations of IL-24B4 and IL-24B4 Y204NBY without and with UV irradiation (Supplementary Figure S12B). The more IL-24B4 was added to the cells the more pSTAT3 was obtained until saturating levels were reached. A non-cooperative Hill model was fitted to the dose-response curves (Supplementary Figure S12C) and the resulting half-maximal effective concentration (EC_{50}) values are reported in Table 1. The EC_{50} of caged IL-24B4



Y204NBY was ~ 3 times larger than that of parental IL-24B4 but similar to the decayed counterpart considering experimental uncertainties. At concentrations higher than EC_{50} , all four IL-24B4 variants behaved similarly as expected (Figure 6). Only at concentrations near EC_{50} , the caged Y204NBY variant signaled less than the parental protein, and UV irradiation increased pSTAT3 levels but without reaching the amounts observed for the parental protein (Figure 6). We conclude that IL-24B4 Y204NBY is another promising candidate for future studies.

4 Discussion

Since the interaction between human interleukin-24 and its receptor IL-20R2 triggers the JAK/STAT signaling pathway, and acts on angiogenesis and cell proliferation, we are interested in controlling this interleukin-receptor system. In this work, we have genetically encoded in each binding partner *ortho*-nitrobenzyl-tyrosine, an unnatural photocaged analogue of tyrosine, to photoregulate IL-24/IL-20R2 interactions. Our main finding is that IL-20R2D Y70NBY and, to a lesser extent, IL-24B4 Y204 hamper complex formation and subsequent activation of

the STAT3 cascade. Mild UV irradiation partially restores heteroassociation and cell signaling.

4.1 Photocontrolling protein-protein interactions involving interleukins with ncAA

Protein-protein interactions (PPI) lie at the heart of many biological processes ranging from metabolic pathways to signaling cascades. Moreover, abnormal PPI are associated with various diseases, including cancer (Lu et al., 2020). Therefore, the modulation of PPI is important not only for basic research but also for the development of new drugs. Research related to the control of PPI has focused primarily on antigen-antibody interactions (Lucchi et al., 2021).

Among several approaches to photocontrol PPI (Carrasco-López et al., 2020; Gil et al., 2020), genetically encoded photo-responsive non-canonical amino acids offer several advantages: i) a precise level of temporal and spatial control over protein binding; ii) the small size of the modification (typical caging groups are less than 500 Da in size); and iii) the possibility of generating the modified protein directly by the biosynthetic machinery within the cell (Courtney and Deiters, 2018). However, there are also challenges associated with their use. First, the suppression efficiencies are variable, context-dependent and may result in severe drops of protein yields, particularly in the case of multi-site incorporation (Amiram et al., 2015). Second, the introduction of non-native side-chains may affect protein folding, stability, binding and dynamics (Rogers et al., 2018; Kesgin-Schaefer et al., 2019; Chaudhari et al., 2023). Third, due to the usually large interfacial areas that are stabilized by multiple non-covalent interactions (hydrogen bonds, hydrophobic interactions, etc.), targeting PPI with single-residue perturbations is not straightforward, typically requiring extensive screening campaigns (Bridge et al., 2019; O'Shea et al., 2022). This contrasts with the photocontrol of interactions between proteins and small ligands or substrates, which typically depend on one or a reduced number of residues closely localized in space, e.g., an enzyme's catalytic site (Nguyen et al., 2014; Zhou et al., 2020).

The interleukin 10 family, which includes IL-24, plays important roles in cell differentiation, apoptosis, and tumor growth, and thus represents an attractive therapeutic target (Chada et al., 2005; Cunningham et al., 2005). Disrupted cytokine-receptor signaling and JAK/STAT phosphorylation cascade may lead to a variety of diseases, such as skin conditions, cancers, and disorders affecting the immune system (Aaronson and Horvath, 2002; Hu et al., 2021). Members of the IL-10 are also involved in host-pathogen interactions, wound healing, and have potential in the treatment of respiratory, inflammatory and autoimmune diseases (Iyer and Cheng, 2012; Traupe, 2017; Liu et al., 2023; Reis et al., 2023). To the best of our knowledge, there is only one example of inhibition of the interaction between IL-24 and its receptors, which is based on antibody binding to IL-20R1 and IL-22R1 (Chada et al., 2004). The light-dependent assembly between IL-24 and its receptors may be an alternative route toward the development of a new therapeutic strategy to combat such diseases and conditions.

4.2 Heterologous expression of interleukins and receptors containing ncAA

To obtain IL-24, the receptor IL-20R2, and their photoactivatable variants, we combine canonical and ncAA mutagenesis.

Depending on protein sequence and need for post-translational modifications, obtaining sufficient yields of stable proteins can be a serious bottleneck. Obtaining “difficult” proteins requires multiple expression attempts (Peleg and Unger, 2012) or cumbersome and cost-ineffective expression systems (Hopkins et al., 2010). On top of these effects, the incorporation of ncAA may cause further decreases in protein yields and alter protein stability or function in ways that are difficult to predict (Chaudhari et al., 2023). Indeed, all our proteins of interest are naturally N-glycosylated and bear disulfide bonds, which make them hard to produce in traditional *E. coli* expression systems. Production of wild-type IL-24 in high quantities is only possible with the co-expression of its interacting partner (Lubkowski et al., 2018). Overall, protein production renders the IL-24 signaling complex a very challenging target for (photo) xenoprotein engineering. Therefore, we adopted an alternative strategy to obtain active proteins while avoiding long and severe optimization trials based on structure- and sequence-based protein design using the 20 canonical amino acids (Goldenzweig et al., 2016). Such a bioinformatics design optimizes the amino acid sequence of IL-24 and IL-20R2 to stabilize their expression in bacteria, rendering them more suitable for functional studies. Both engineered variants, named IL-24B4 and IL-20R2D, are biologically active and support signal transduction through the JAK/STAT pathway.

Next, the light-inducible variants are made by adding a photocaged amino acid, NBY, into IL-24B4 and IL-20R2D by genetic code expansion technology. We looked for candidate residues in both interacting partners and found three target tyrosines. The single tyrosine mutant of IL-24B4, Y204NBY, did not express at sufficient and cost-effective levels for subsequent biophysical investigations by MST. Thus, IL-24B4 Y204NBY could only be studied in cell assays, which require relatively low amounts of material. In the case of the receptor, all ncAA variants including two single mutants and one double mutant were purified at the required amounts for all three methods employed in this work (MST, yeast display and human cell assays). NBY did not substantially affect protein conformation but it had a negative impact on protein's thermal stability, particularly in the case of the double mutant.

4.3 Monitoring IL-24 – IL-20R2 interactions by biophysical methods and cell assays

IL-24B4 Y204NBY exhibits minimal albeit reproducible differences in cell signaling assays when comparing parental and caged variants.

We find one clear “hit” confirmed by both MST and yeast display: IL-20R2D Y70NBY shows low affinity for IL-24B4 and the interaction can be activated by UV light to reach native-like levels. This in line with previous studies reporting that a single NBY residue is sufficient to diminish PPI (Bridge et al., 2019;

Jedlitzke et al., 2019; Joest et al., 2021; Jedlitzke and Mootz, 2022). Here we used an innovative approach of yeast surface display to gain better control over our experiment and see the residual concentrations. We mixed yeasts expressing IL-24B4 with a low amount of yeasts expressing its affinity matured counterpart, which has more than two orders of magnitude higher affinity (IL-24S). The analysis showed detectable but low, below saturation, binding signals of IL-24S with IL-20R2D Y70NBY (Supplementary Figure S10). Given the difference in binding affinity, the concentration of binding-capable ligands was below 1% of total concentration.

Unfortunately, due to limitations in our signaling assays, including both the choice of synthetase and the background signaling effect of IL-20R2D, we could not test the effect of IL-20R2D Y70NBY in human cells. It was previously reported that dimerization of IFN α 2 is sufficient for induction of interferon-regulated genes but not for full activity (Pattyn et al., 1999). Here, we show that purified IL-20R2D added to the cells can trigger signaling albeit at high concentrations. This problem may be partially circumvented by using general reporter cell lines and/or reporter assays based on gene expression (Mock et al., 2020; Cho et al., 2023) with further research needed. As an alternative to the genetic encoding of NBY in sequence-optimized truncated receptors recombinantly expressed in *E. coli*, one could biosynthetically incorporate photocaged tyrosines in full-length IL-20R2 receptors expressed in HeLa cells by using an orthogonal translation system derived from *Methanosarcina barkeri* (Arbely et al., 2012).

The more effective blocking of heterocomplex assembly with Y70NBY compared to Y204NBY can be rationalized in structural terms: Y70 of IL-20R2 is found at the core of the IL-24/IL-20R2 interface while Y204 of IL-24 is located in a flexible region at one edge of the interaction surface. We notice that the use of NBY to photocontrol IL-24-dependent PPI necessitates further improvements. Y70NBY and particularly Y204NBY do not fully prevent complex assembly, but rather shift the binding equilibrium towards higher concentrations (lower affinities). IL-24/IL-20R2 complex may still be formed and signal transduction started if large concentrations of photocaged binding partners are employed. In principle, substantial affinity differences between native and caged variants are preferred. The dynamical range is much narrower for IL-24B4 Y204NBY (3-fold difference in EC_{50}) than for IL-20R2D Y70NBY (at least 300-fold difference in K_d according to both MST and yeast display). Larger dynamical ranges could be achieved through the installment of photocages on each partner, for instance Y70NBY in IL-20R2D and Y204NBY in IL-24B4. Although residues at the binding interface seem the natural choice for replacement, other distant sites may also influence PPI through allosteric communication, which has been reported for some interleukins (Bowman and Geissler, 2012; De Paula et al., 2020).

Strikingly, the two single mutants (IL-20R2 Y70 NBY and IL-20R2D Y74NBY) and the double mutant (IL-20R2D Y70NBY/Y74NBY) show clear signs of non-additive effects on protein binding. Y70NBY inhibits association with IL-24B4 while Y74NBY and Y70NBY/Y74NBY do not significantly alter protein-protein interactions. Non-additive interactions between mutations (epistasis) are common in proteins and play a crucial

role in protein engineering (Reetz, 2013). There is no clear picture of the mechanisms that cause epistasis, although effects on protein stability, conformation and dynamics have all been invoked (Starr and Thornton, 2016). Our CD and thermal denaturing experiments suggest minimal differences among the mutants. All variants show similar secondary structure content. The melting temperature of the mutants is reduced with respect to the parental proteins and the T_m of the double mutant Y70NBY/Y74NBY is further reduced thus suggesting additive effects. Therefore, protein structure and thermal stability do not seem to be the major causes of the negative (antagonistic/deleterious) epistasis observed when combining Y70NBY and Y74NBY mutations. We hypothesize that the two mutations together alter IL-20R2D structural dynamics, a phenomenon that has been described for other proteins (Wagner et al., 1995; Acevedo-Rocha et al., 2021).

Interestingly, we found evidence that in some cases UV-irradiation alone reduced the binding affinity. Although we did not detect clear signs of photoinduced chemical modifications, e.g., photo-oxidation (the masses of IL-24B4 and IL-20R2D before/after UV are virtually identical), the affinity between IL-24B4 and IL-20R2 was reduced 6-fold (MST-based experiments) or 2-fold (yeast surface display experiments). This result implies that our UV irradiation protocol damages the proteins to a certain extent. We suggest that by carefully controlling the illumination conditions (wavelength, power density and time) such an undesired effect may be eliminated. Moreover, UV light can potentially damage nucleic acids depending on the dose used and it does not penetrate deeply into tissues (Kielbassa et al., 1997). Consequently, for *in vivo* applications requiring native-like binding affinities it would be beneficial to use other photocages. Suitable candidates would be the nitropiperonyl moiety, which features longer absorption wavelengths (Gautier et al., 2010; Luo et al., 2017) or coumarins, which can be activated by two-photon excitation in the near infrared range (Luo et al., 2014). Apart from photocaged ncAAs, which enable unidirectional OFF-to-ON switch of protein binding by light, one could use other ncAA. For instance, in applications requiring bidirectional photocontrol of protein binding (Jankovic et al., 2019; Myrhammar et al., 2020; Zhang et al., 2022), photoswitchable ncAA like azobenzene-phenylalanine constitute an excellent alternative (Bose et al., 2005; Luo et al., 2018; Israeli et al., 2021).

Overall, our results suggest the feasibility of regulating interactions involving protein partners from the interleukin family, IL-24 and IL-20R2, with UV light and photocaged tyrosines introduced at certain residue positions. In the future, other cytokine/receptor pairs and non-antibody-based protein scaffolds (Kolářová et al., 2021; Pham et al., 2021; Huličák et al., 2023) may benefit from a similar approach and offer excellent interaction control.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://doi.org/10.5281/zenodo.7877998>.

Author contributions

Conceptualization, GF; methodology and investigation, PP and JZ; resources, LK and GF; writing—original draft preparation, PP and GF; writing—review and editing, PP, JZ, BS and GF; project administration and supervision, GF; funding acquisition, JZ and BS. All authors contributed to the article and approved the submitted version.

Funding

The work was supported by the project ELIBIO: Structural dynamics of biomolecular systems (CZ.02.1.01/0.0/0.0/15_003/0000447) from the European Regional Development Fund and the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic (CR). The Institute of Biotechnology of the Czech Academy of Sciences acknowledges the institutional grant RVO 86652036. We acknowledge CF Biophysic, CF SMS of CIISB, Instruct-CZ Centre, supported by MEYS CR (LM2023042) and European Regional Development Fund-Project “UP CIISB” (No. CZ.02.1.01/0.0/0.0/18_046/0015974). We acknowledge Charles University grant PRIMUS no. 23/MED/002. This paper was funded by the European Union - Next Generation EU (National Institute for Cancer Research, programme EXCELES, Project No. LX22NPO5102).

Acknowledgments

We thank Petr Pompach and Pavla Vaňková for the mass spectrometry measurements. We are also grateful to Tania Charnavets and Volha Dzmitruk for assistance with MST analysis.

We thank Aditi Chatterjee for calibrating the gel filtration column.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2023.1214235/full#supplementary-material>

