T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

DEVELOPMENT OF ANTI-CANCER ANTIBODY FRAGMENTS WITH IMPROVED CHARACTERISTICS THROUGH ANTIBODY ENGINEERING APPROACHES

MERVE ARSLAN

DEPARTMENT OF GENOMICS AND MOLECULAR BIOTECHNOLOGY

DOCTOR OF PHILOSOPHY THESIS

İZMİR-2023

THESIS CODE: DEU.HSI.PhD-2017850058

T.C.
DOKUZ EYLUL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

DEVELOPMENT OF ANTI-CANCER ANTIBODY FRAGMENTS WITH IMPROVED CHARACTERISTICS THROUGH ANTIBODY ENGINEERING APPROACHES

DEPARTMENT OF GENOMICS AND MOLECULAR BIOTECHNOLOGY

DOCTOR OF PHILOSOPHY THESIS

MERVE ARSLAN

Advisor: Assos. Prof. Sinan Güven

Co-advisor: Sibel Kalyoncu, PhD

This thesis was supported by YÖK 100/2000 Doctoral Scholarship, TÜBİTAK BİDEB 2211-A National Doctoral Scholarship, TÜBİTAK BIDEB 2214-A International Research Doctoral Fellowship, and EMBO Scientific Exchange Grant.

THESIS CODE: DEU.HSI.PhD-2017850058

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
İZMİR ULUSLARARASI BİYOTIP VE GENOM ENSTİTÜSÜ

DOKTORA TEZ SAVUNMA SINAVI TUTANAĞI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı Moleküler Biyoloji ve Genetik Bütünleşik Doktora Programı 2017850058 numaralı öğrencisi Merve Arslan, "Development of Anti-Cancer Antibody Fragments with Improved Characteristics Through Antibody Engineering Approaches" konulu Doktora tezinin 05.06.2023 tarihinde yapılan savunma sınavı sonucunda başarılı / ~~başarısız~~ olmuştur.

BAŞKAN
Doç. Dr. Sinan Güven
Dokuz Eylül Üniversitesi

ÜYE
Dr. Öğr. Üyesi Hani Alotaibi
Dokuz Eylül Üniversitesi

ÜYE
Prof. Dr. Gülgün Oktay
Dokuz Eylül Üniversitesi

ÜYE
Prof. Dr. Mehmet İnan
Akdeniz Universitesi

ÜYE
Doç. Dr. Burcu Kaplan Türköz
Ege Universitesi

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Association rate constant	K_{on}
Basic Local Alignment Search Tool	BLAST
Binding free energy	ΔG
Bio-layer interferometry	BLI
Buried surface area	BSA
Complementarity determining region	CDR
Degrees Celsius	° C
Deoxyribonucleic acid	DNA
Differential Scanning Fluorimetry	DSF
Dissociation rate constant	K_{off}
dissociation constant	K_D
Dual targeting Fab	DutaFab
Enzyme-linked immunosorbent assay	ELISA
Escherichia coli	*E.coli*
Ethylenediaminetetraacetic acid	EDTA
Fluorescence-Activated Cell Sorting	FACS
Fragment antigen binding	Fab
Fragment crystallizable	Fc
Framework	FW
Glycine-serine	GS
Heavy chains	HC
Heavy constant domains	C_H1, C_H2, C_H3

Heavy variable domain	V_H
Heavy Vernier	HV
HEPES buffered saline	HBS
High-performance liquid chromatography	HPLC
Human Epidermal Growth Factor Receptor 2	HER2
Imidazole	ImH
Immune Epitope Database	IEDB
Isothermal titration calorimetry	ITC
Kelvin	K
Kilodalton	kDa
Kilovolt	kV
Light chain	LC
Light constant domain	C_L
Light variable domain	V_L
Light Vernier	LV
Liter	L
Luria-Bertani	LB
Microscale thermophoresis	MST
Micro	µ
Millisecond	msec
Molar	M
Molecular dynamics	MD
Molecular weight cut-off	MWCO
Nuclear magnetic resonance	NMR
Optical density measure at 600 nm	OD_{600}

Phosphate-buffered saline	PBS
Pichia pastoris	*P.pastoris*
Placental growth factor	PIGF
Polymerase chain reaction	PCR
Programmed Death-Ligand 1	PD-L1
Protein Data Bank	PDB
PDBe Protein Interfaces, Surfaces and Assemblies	PDBePISA
protein with internal repeats of *P.pastoris*	PpPIR1
Real-time polymerase chain reaction	*RT-PCR*
Refractive index detector	RID
Response units	RU
Root mean squared deviation	RMSD
Root mean square fluctuation	RMSF
Rotor per minute	RPM
Saccharomyces cerevisiae	*S. cerevisiae*
S. cerevisiae cell wall agglutinin protein 1	Sag1
Single-chain antibody fragments	scFvs
Single domain antibody	sdAb
Size exclusion chromatography	SEC
Sodium dodecyl-sulfate polyacrylamide gel electrophoresis	SDS-PAGE
Solvent-accessible surface area	SASA
Subintestinal vessel	SIV
Surface plasmon resonance	SPR
Thermal melting point	T_m

T cell receptor	TCR
variable regions	Fv
Vascular Endothelial Growth Factor	VEGF
Vascular Endothelial Growth Factor Receptor	VEGFR
Weight/volume	w/v
Western blot	WB
Wild type	WT

ACKNOWLEDGEMENT

This is the end of my eleven years of Izmir era, and the end of almost six years of my Ph.D. adventure. I have tried to do my best to go one step further and thanks to everyone who has been with me during these years. My special thanks;

I would like to thank the Council of Higher Education of Turkey for their financial support under the program YÖK 100/2000 Ph.D. Scholarship, The Scientific and Technological Research Institution of Turkey (TÜBİTAK) for their financial support under the programs of TÜBİTAK BİDEB 2211-A National Doctoral Scholarship and TÜBİTAK BİDEB 2214-A International Research Fellowship, European Molecular Biology Organization (EMBO) for their financial support under the program of Scientific Exchange Grant, DEU-IBG Institute and IBG Center and each laboratory for their support, for sharing anything needed that made me able to continue conducting my studies. I also would like to thank IBG Optical Imaging, IBG Flow Cytometry, IBG Zebrafish, and IBG-Pharma facilities for their contributions to my research.

I would like to thank my institute advisor, Dr. Sinan Güven, his valuable support is much appreciated during my doctorate program. I would like to thank my committee members, Dr. Hani Alotaibi, Dr. Gülgün Oktay, and Dr. Burcu Kaplan Türköz for their input to this dissertation. I would like to thank Dr. Mehmet İnan for both being my committee member and for his guidance throughout my thesis.

I would like to thank our collaborators for their valuable contribution to my work, Dr. Gülçin Çakan-Akdoğan and members of Çakan Lab, for their unlimited sharing, Dr. Seyit Kale for valuable discussions, formal member of Kale Lab, Tuğçe Uluçay, for her fantastic academic partnership, Dr. Nico Callewaert for his great support to my research and all the opportunities he has provided to me during my visit, VIB-UGent Flow Facility, members of Callewaert Lab for being embracive to me, and Hannah Eeckhaut for her help during my visiting period.

I would like to thank all the current and former members of Kalyoncu Lab, Nazlı Eda Kaleli, Dilara Karadağ, Murat Karadağ, Ayşe Ergündoğan, Cansu Ergün and members of IBG-Pharma, Hüseyin Akıntürk, and Şeyda Güllü for their help in any case.

I would like to express my deepest gratitude to Mesut Soylu and the family at Axioms for unlimited support regardless of the subject, many precious brainstorming, proofreading of hundreds of pages, and for possible many works to do together.

I owe my most enormous thanks to my mentor Dr. Sibel Kalyoncu, she created such a supportive environment and shared every piece of knowledge with me no matter what. She made me feel like a princess in the harsh conditions of a Ph.D. I will always be grateful for the effort she puts on me. I am honored to make a lifetime collaboration with such an inspiring scientist.

My lab mates, my feelings for you guys are beyond appreciation. I thank each and every one of you for being my friend, for sharing my ignorance. I cannot place you in any group or refer you as current/formal as this can change quickly, who knows! Hakan, it was a pleasure to be lunch buddy and lately thesis buddy. You are going to suffer from the PhD program soon, right?! Aslı and Saniye, spending time with you is always more than sharing the lab. Thanks for being my fellas in any Balkan song, in any nonsense comedy. Ceren, thank you for being there for me anytime. Your help with my studies is much appreciated. I love that we worked in perfect harmony in the lab, without distinction between your or my work. It is priceless to have you around. My -current - former - to be current soon- lab mates, my roommates, Erhan and Semiramis, it is just unbelievable how much we share during the years. Erhan thank you very much for being my writing buddy, for being such a virgo in my life. Semiramis, ah Semsem, you are one extraordinary character that must be involved in my life. I would never have guessed that we would all end up in Belgium together when we first met. I hope we will have many full-of-surprise moments in the future.

My sincere thanks go to my lifetime friends, my sisters, İdil and Melike, thanks for your encouragement and for being my antidepressant anytime I needed it. I believe in us that we will normalize all kinds of nonsense to make it easier for each other. What could go wrong, right?!

Finally, I would like to thank my family, my dad Sadi, my mom Seher, my brother Onur and my sister-in-law Ece for their love and support.
I am so glad that My PhD journey is full of good memories and, I was able to do what I wanted to do. I wish my beloved nieces the same feelings no matter what they do in their future. I dedicate this thesis to my baby girls, to Burçak, and to Açelya…

DEVELOPMENT OF ANTI-CANCER ANTIBODY FRAGMENTS WITH IMPROVED CHARACTERISTICS THROUGH ANTIBODY ENGINEERING APPROACHES

Merve Arslan,
Dokuz Eylül University Izmir International Biomedicine and Genome Institute,

ABSTRACT

Cancer is a complicated disease influenced by various factors, cells, and signaling pathways. Vascular Endothelial Growth Factor (VEGF) induces angiogenesis and can be found in several cancer types. Since VEGF acts as immunosuppressive in tumor environment, dual-blockade of angiogenesis and immune checkpoints such as Programmed Death-Ligand 1 (PD-L1) becomes a successful combinatorial cancer therapy.

Antibody formats, such as single-chain antibody fragments (scFvs), are detailly-studied for targeted therapy approaches. Yet, they require antibody engineering to enhance developability properties such as affinity, stability, and specificity. One of the important regions for antibody engineering is the Vernier zone regions. Vernier zone is selected residues within antibody's framework and are capable of affecting antigen binding. The impact of Vernier zone on other antibody characteristics like stability and specificity has yet to be fully understood.

Controlled multi-targeting combinatorial therapies show outstanding success in cancer over mono-specific targeting. One controlled multi-specific format is dual-specifics which hold the advantage of having two antigen recognition capacities for the size of one.

Antibody formats, both in constructional and targeting manners, against VEGF and/or PD-L1 remain uninvestigated in terms of the recent antibody engineering applications. In this thesis, I aimed to develop improved anti-VEGF scFv fragments for cancer therapies. First, I generated anti-VEGF scFv with improved affinity-stability. Further, I established a novel strategy to modulate the mono-specificity of generated anti-VEGF scFv to dual-specificity against PD-L1 to increase the therapeutic efficacy. Many valuable implications on the importance of Vernier zone regions on affinity, stability, and specificity have been highlighted within thesis chapters.

Keywords: Anti-Vascular Endothelial Growth Factor, single-chain antibody fragments, antibody engineering, antibody developability, Vernier zones, dual-specificity

ANTİKOR MÜHENDİSLİĞİ YAKLAŞIMLARI İLE İYİLEŞTİRİLMİŞ ÖZELLİKLERE SAHİP ANTİ-KANSER ANTİKOR FRAGMANLARININ GELİŞTİRİLMESİ

Merve Arslan,
Dokuz Eylül Universitesi Izmir Uluslararası Biyotıp ve Genom Enstitüsü

ÖZET

Kanser, çeşitli faktörlerin, hücrelerin ve sinyal yolaklarının katılımıyla ortaya çıkabilen karmaşık bir hastalıktır. Vasküler Endotel Büyüme Faktörü (VEGF) tümör anjiyogenezini uyarır ve çeşitli kanser türlerinde bulunabilir. Ayrıca VEGF immünosüpresif görevi görmektedir, buna bağlı olarak anjiyogenez ve immün kontrol noktalarının, Programlanmış Ölüm-Ligand 1 (PD-L1) gibi, ikili blokajı başarılı kombinatoryal kanser tedavisi sunmaktadır.

Tek zincirli antikor fragmentleri (scFv) gibi antikor formatları, hedefe yönelik tedavi yaklaşımları için çokça çalışılmaktadır. Yine de, afinite, stabilite ve spesifilik gibi geliştirilebilirlik özelliklerini iyileştirmek için antikor mühendisliği gerekmektedir. Antikor mühendisliği için önemli bölgelerden biri Vernier bölgeleridir. Vernier bölgeleri, bir antikorun çerçevesi (framework) boyunca seçilmiş rezidülerdir ve antijen bağlanmasını etkileme yeteneğine sahiptir. Spesifiklik ve stabilite gibi diğer antikor özellikleri üzerinde Vernier bölgesinin rolleri açıklanmayı beklemektedir.

Kanserde monospesifik hedefleme ile başarılı bir tedavi elde etmek zordur ve çoklu hedeflemeli tedaviler kanserde üstün başarı göstermektedir. Bu yaklaşımlardan biri, aynı antijen bağlama bölgesi ile iki farklı antijene bağlanabilen dual-spesifik antikorlardır. Dual-spesifik antikorlar, bir antikor sekansıyla iki antijen tanıma kapasitesiyle avantajlıdır.

VEGF ve/veya PD-L1'e karşı farklı yapısal ve hedefleme antikor formatları, son antikor mühendisliği uygulamaları açısından henüz araştırılmamıştır. Bu tezde kanser tedavileri için iyileştirilmiş özelliklerde scFv fragmentleri geliştirilmesi amaçlanmıştır. Bu amaç için, afinite-stabilite özellikleri iyileştirilmiş anti-VEGF scFv geliştirilmiştir. Devamında, geliştirilen anti-VEGF scFv'nin terapötik etkinliğini artırmak için, Vernier bölgeri üzerinden antikora yeni spesifiklik kazandırma üzerine özgün bir strateji geliştirilmiş, anti-VEGF scFv'nin monospesifik özelliği PD-L1'e karşı dual-spesifik özelliğe çevrilmiştir. Tez bölümlerinde, Vernier bölgelerinin afinite, stabilite ve özgüllük üzerindeki önemine ilişkin değerli çıkarımlar tartışılmıştır.

Anahtar Kelimeler : Vasküler Endotel Büyüme Faktörü, Tek zincirli antikor fragmentleri, antikor mühendisliği, antikor geliştirilebilirliği, Venier Bölgeleri, dual-spesifiklik

1. INTRODUCTION AND AIM

1.1. Antibody structure and function

Antibodies are glycoproteins that include two identical light (LC) and heavy chains (HC). Associated light and heavy chain forms a heterodimer through disulfide bonds. Pairing with another identical heterodimer through disulfide bonds constructs the intact antibody structure [1]. In the light chain, there are a constant domain (C_L) and a variable domain (V_L). There are two classes, V□ and V□□□ for the human light chain [2]. The heavy chain includes a variable number of domains due to its according to isotypes. IgA, IgD, and IgG have three constant domains (C_H1, C_H2, C_H3) and a variable domain (V_H) while IgE and IgM contain an extra domain (C_H4). IgG is the superabundant antibody in the circulating system that takes around 75 % of the human immunoglobulins [3]. All the domains of antibody heavy and light chains share the same structural conformation called "immunoglobulin fold". This fold is mainly specified with anti-parallel β-sheets and loops [4].

Antibody structure at its simplest can be divided into two; a region for antigen-binding, called fragment antigen binding (Fab), and the receptor binding region, called fragment crystallizable (Fc). Fc consists of the hinge region, C_H2, and C_H3 domains. Two C_H3 domains pack each other, providing stability of the Fc region and overall antibody. Two C_H2 domains have little or no contact between [5]. However, domains have a glycosylation site at residue N297 that different oligosaccharide motifs alter the effector immune functions [6]. The interface between C_H2 and C_H3 serves as an epitope for Protein A, Protein G, and Fc receptors, effects the C_H2 orientation, thus the Fc receptor binding [7]. The hinge region between C_H1 and C_H2 has three sub-regions; the N-terminal upper hinge enables the Fabs to move in a significant level of conformational flexibility abiding by Fc, the core hinge consists of several cysteine residues therefore, disulfide bonds that stabilize heavy chain dimerization, C-terminal lower hinge enables the Fc to rotate relative to Fab and some residues have potential to be included in Fc receptor binding [8].

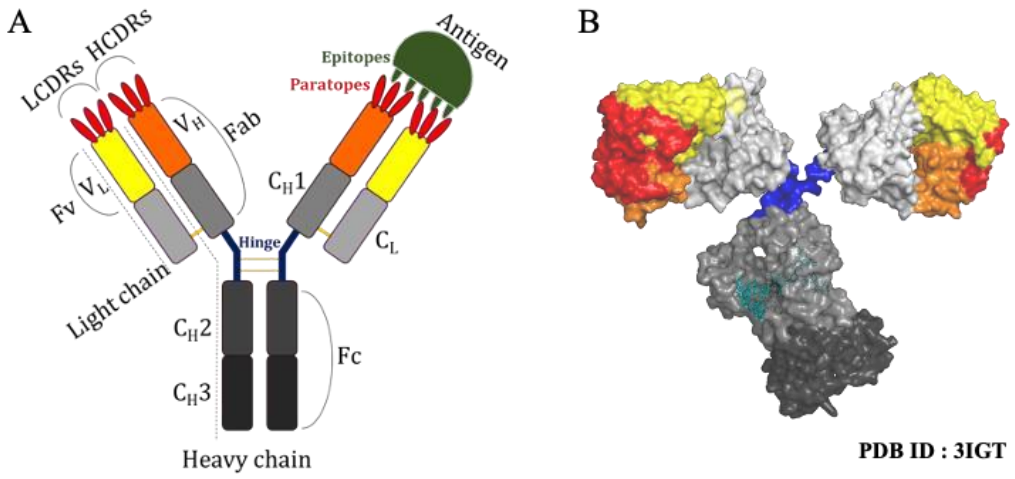


Figure 1. 1 Antibody (IgG) structure

A) Schematic and B) Crystal structure of the IgG. Yellow lines indicate disulfide bonds. The glycans are shown in cyan in the crystal structure.

The Fab region comprises paired variable regions, V_H and V_L, and paired constant regions, C_H1 and C_L. Constant domains pack closely and utilize a disulfide bond while variable regions pack through only hydrophobic interactions. The orientation between variable and constant domains called elbow angle affects the overall flexibility and the stability of the Fab [9]. Each variable region consists of nine □-sheets, packed together through non-covalent interactions and one disulfide bond [4].

The antigen-binding site, paratope, within variable regions (Fv) is formed by six loops called complementarity determining regions (CDRs) on variable chains (LCDR1-3 and HCDR1-3). These loops are intrinsically hypervariable, in both the primary sequences and the length, naturally originate from somatic hyper-mutations in mature B cells, providing diversity and required specificity [10]. CDRs have a unique “canonical structure” defined by their loop conformation, loop length, conserved amino acids within the loop, and framework nearby the CDRs [11]. Unique canonical structure affects the overall features of antibody-antigen interaction.

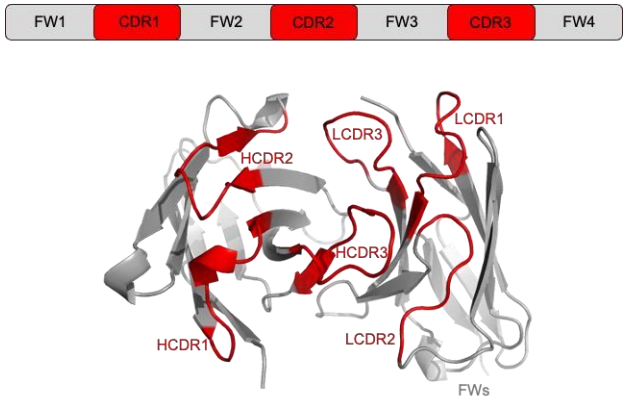


Figure 1. 2 Variable regions

Schematic (A) and (B) crystal structure of variable regions.

Antibodies' target protein is called an antigen, the regions within the antigen recognized by antibodies are referred to as epitopes [12]. Paratopes can recognize (i) linear or continuous epitopes where the residues within the epitope are continuous in primary sequence, (ii) conformational or discontinuous epitopes where the residues within the epitope are from different conformational regions, and come together due to secondary structure, and (iii) hybrid epitopes where the linear epitopes are forms a conformational epitope due to secondary structure and can be recognized in both ways [13].

CDRs are separated by highly conserved non-CDR regions called frameworks (FWs) that constitute a core β-sheet structure and display the CDRs on the tip of the variable domains (Figure 1.2) [14]. FWs are essential for antibody folding which can alter the biophysical and biochemical properties of the antibody [15, 16]. The most dramatic effect of the FWs on antibodies can be observed during the humanization approaches of non-human antibodies. Humanization is an antibody engineering application to reduce the immunogenicity of the antibody produced in non-human species such as mouse [17], rabbit [18], and chicken [19]. CDR grafting is the most common approach for humanization that CDRs from non-human parental antibodies are transferred within the human frameworks [20]. Although human frameworks are chosen based on the highest homology with the parental antibody frameworks [21, 22], the antibody affinity is reduced in most of the cases [23-25]. Thus, CDR grafting is generally followed by back mutations of some residues to parental antibody residues to re-gain antibody affinity [24].

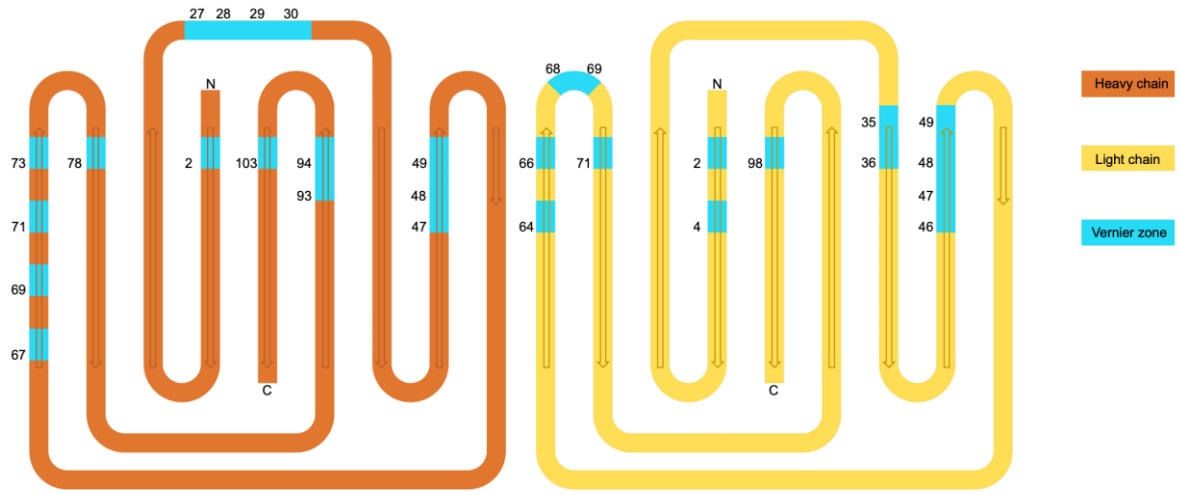


Figure 1. 3 Vernier zone residues on variable frameworks

In 1988, it is shown that some residues that are back mutated are really important in the packing and loop structure of the CDRs [26]. Later they are called Vernier zone residues [27], as they are located within the FW regions and play a foundational role in the CDR structure and antigen binding. Vernier zone includes 30 residues, 16 in the heavy variable region and 14 in the light variable region (Figure 1.3). Conformationally close Vernier zone residues pack and act together in the interaction network to support CDRs to have proper canonical structure for antigen binding. Thus, modification of Vernier residues affects the CDR reshaping, therefore humanization applications rely on retaining of these important residues [28]. However, it is also shown that it is not necessary to back-mutate all the Vernier zone residues to regain antibody affinity. Some of the residues contribute drastically to affinity based on the closeness to the main CDR in the antigen interaction [29]. For example, Bevacizumab (Avastin), an approved anti-VEGF antibody, developed in 1997 by humanization including CDR grafting and back mutations of total of 7 residues; 1 Vernier zone residue in light chain, 5 Vernier zone residues in heavy chain and 1 FW residue in heavy chain (Figure 1.4) [17].

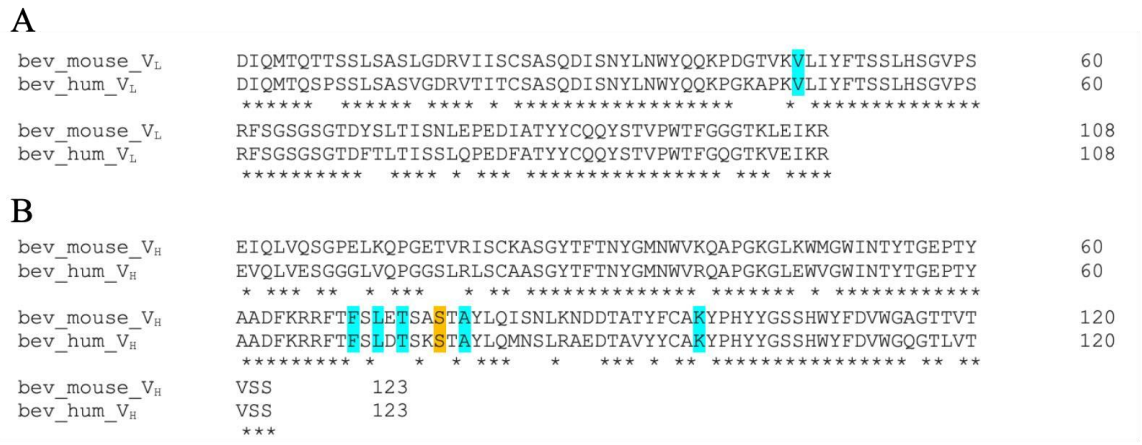


Figure 1. 4 Humanization of Bevacizumab

Back mutated residues of (A) V_L and (B) V_H. Vernier zone residues and framework residues are colored cyan and orange, respectively.

1.2.Antibody affinity and specificity

Antibody-antigen interactions require binding to its target with high specificity and with a sufficiently high affinity as characterized by nanomolar to picomolar dissociation constants, K_D. The sequence and length diversity of the CDRs due to the higher number of possibilities of V(D)J combinations of genes mainly determine the antibody affinity and stability. The CDR3s of heavy chain and light chain are the most variable ones, while the other CDRs are only encoded by the V-gene segment, resulting in a decreased number variation compared to CDR3s [30]. Besides, HCDR3 exhibits substantial length variation among natural antibodies, making it the most diverse and crucial region for determining affinity [30, 31]. Furthermore, apart from the residues directly involved in binding, non-CDRs typically play roles in the "canonical structure" of CDR loops, which is essential for antigen recognition, as explained in section 1.1 [32].

The parts of an antibody and an antigen that form the binding interface are called paratopes and epitopes, respectively, and the binding strength is called antigen affinity. Paratopes are typically enriched in aromatic side chains (phenylalanine, tyrosine, tryptophan) which dominate interfacial contacts, while smaller and hydrophilic amino acids (serine, threonine, asparagine, aspartate) surround the aromatic residues and provide hydrogen bonding between the epitope

and the paratope [33, 34]. Thanks to its unique physicochemical nature, tyrosine provides two advantages for favorable contacts: (i) amphipathicity, i.e., the ability to tolerate the change from hydrophilic to hydrophobic upon protein binding [35], and (ii) the capacity of forming variable contacts, including nonpolar, hydrogen-bonding and π-interactions. Furthermore, tyrosine is relatively inflexible, which limits the loss of conformational entropy upon binding, thereby contributing to a higher specificity [36].

The binding strength of an antibody to its target can be expressed via a dissociation constant (K_D) or a binding free energy (ΔG). K_D is the protein concentration (antigen or antibody, whichever is limiting) at which the number of bound complexes in solution is equal to those unbound. This is the point of kinetic equilibrium at which for every cognate protein pair that dissociates, a new one forms right away. A low K_D value implies that the antibody-antigen pair has either a high rate of association (meaning they can rapidly diffuse toward each other) or a low rate of dissociation (meaning the bound complex has a high lifetime), or a combination of both. Because the rates of association show little variation among interacting protein pairs, it is usually the rate of dissociation, that is the primary determinant of K_D [37].

Dissociation constants are expressed in concentration units, and nano- to picomolar values are common among high-efficacy antibodies [38]. The experimental techniques for measuring K_D involve surface plasmon resonance (SPR) [39], isothermal titration calorimetry (ITC) [40], nuclear magnetic resonance (NMR) chemical shifts [41], spectroscopic measurements via fluorescent labeling [42], microscale thermophoresis (MST) [43], and the more recently developed bio-layer interferometry (BLI) [44]. Enzyme-linked immunosorbent assay (ELISA) is a relatively faster alternative, while accuracy could be a drawback [45]. There are also *in silico* prediction tools such as Prodigy [46], mCSM-AB2 [47], BeAtMuSiC [48], MutaBind2 [49].

The success of antibodies for its applications lies heavily on the specificity of antibody-antigen interactions. Multi-specificity, also known as cross-reactivity, is a significant phenomenon in the immune recognition [50]. Multi-specificity is typically associated with the potential for self-reactivity and/or autoimmunity. However, it has been shown that it is a conserved event of the immune system [51]. Antibody's ability to acquire multi-specificity through somatic mutations impacts their omnipresence in the repertoire [52, 53]. Natural antibodies exhibit a degree of

inherent specificity, although only a limited number of them possess the ability to interact with multiple antigens with significant affinities [54]. Recent findings suggest that multi-specific antibodies could play a critical role in enhancing the immune system's repertoire [55, 56]. Nevertheless, due to evolutionary pressures, there is a delicate equilibrium between effectively combating pathogens and preventing autoreactivity, even though the presence of multi-specificity carries the risk of triggering autoimmune responses. [57, 58].

The understanding of the specificity of antibodies remains a complicated challenge in the antibody development [59]. Exploring the underlying reasons for mono-specificity, multi-specificity, and non-specificity is highly challenging. While multi-specificity may have some advantages in therapeutic approaches, caution must be exercised to avoid compromising the antibody developability [60]. Studies have demonstrated that approved antibodies tend to exhibit greater specificity compared to those in clinical trials [61] and this discrepancy in specificity could be influenced by various factors, such as the biochemical properties of the amino acids within CDRs [62]. The impact of several properties, including glycosylation, charge, and hydrophobicity on the antibody specificity has been shown specificity [63-66].

1.3.Antibody formats

Therapeutic antibodies are well-established tools in targeted therapy approaches. Numerous antibodies can be found in various *constructional formats* such as Fab, scFv, nanobodies ((V_HH or single domain antibody (sdAb)), and *targeting formats* such as diabodies, bispecific, dual-specific antibodies, continue being developed for their therapeutic efficacies [67]. The most common format for therapeutic antibodies is mono-specific full-length IgG structure. There are currently 158 full-length antibodies, 4 Fab, 1 scFv, 1 sdAb, 1 sdAb-Fc, and 1toxin conjugated disulfide-stabilized Fv antibody fragment (dsFv) approved for mono-specific targeting (extracted from Antibody Society, approved therapeutic monoclonal antibodies, date accessed: 21 April 2023). However, several factors, different cells, and signaling pathways are involved in cancer. Thus, it is difficult to achieve a successful treatment with mono-specific targeting and combinatorial therapies are successful in the cancer [68-70]. Bispecific antibodies, which possess two distinct binding sites, are the most prevalent format for multi-targeting. These antibodies exhibit affinity towards two different antigenic epitopes. The first bispecific antibody is proposed in the 1960s [71, 72], the concept and the approaches to generate bispecific

antibodies are improved with the recent engineering studies [73]. There are currently 10 fulllength antibodies, 1 tandem scFv (diabody with two different scFv), 1 triple sdAb, 1 CrossMab (Fab x Fab-Fab-Fc) and 1 scFv- T cell receptor (TCR) fusion protein approved for bispecific targeting (extracted from Antibody Society, approved therapeutic monoclonal antibodies, date accessed: 21 April 2023). In addition to these, diabody is an antibody format in that the same antibody fragments are fused to increase avidity, accumulated binding strength of the multiple affinities [74, 75]. Another extensively studied format is the dual targeting antibody, which can bind two different antigens using a shared sequence. This approach is established in Fab fragments as an engineering platform, therefore called dual targeting Fab (DutaFab) [76]. The most common constructional and targeting formats are represented in Figure 1.5.

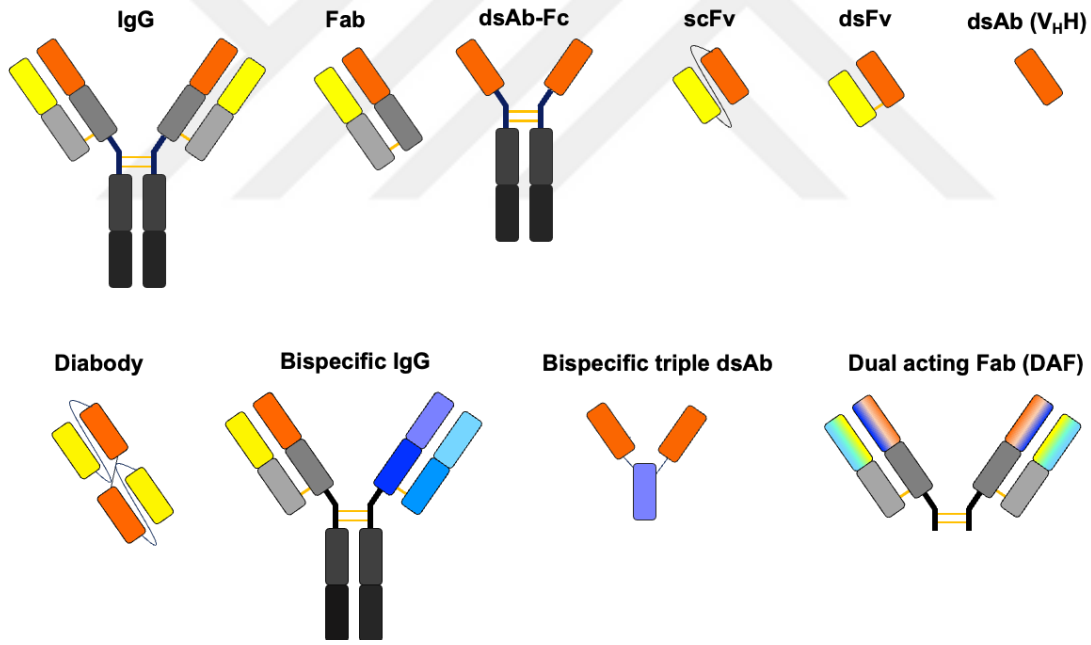


Figure 1. 5 Antibody formats

1.4.Antibody engineering

Since its first literature appearance of the term “antibody” nearly 130 years ago [77], antibodies continue to be studied and developed extensively. The introduction of the hybridoma

technology [78] and the first reported crystal structure in the 1970s [79, 80] pioneered a groundbreaking phase in antibody engineering. The irreplaceable role of antibodies in the adaptive immunity [81, 82] is intrinsically linked to their biopharmaceutical potential, encouraging the antibody developability efforts of the modern pharmaceutical industry and academia. The current market size of the therapeutic was around 3.6 billion Dollars last year, and it is expected to grow annually by 6.4% through 2028 [83].

Discovery is the very first step in the development cycle of an antibody, and it can be immunization-dependent or immunization-independent. Methods of discovery include *in vitro* [84-86] and *in vivo* [87-89] display systems, transgenic animals [90, 91], or human B cell derivation [53, 92, 93]. After the discovery, the antibody undergoes a set of routine engineering steps to improve its biophysical and biochemical properties. These include developability properties which at the discovery stage often don't meet the expectations of a therapeutic profile. Important developability properties are affinity, specificity, stability, and solubility/aggregation and generally, it is challenging to predict these properties that affect the overall antibody success. The challenges associated with developability can arise during the production process due to the intricate and complex nature of the protein. These challenges can lead to decreased antigen binding affinity, immunogenic responses, and waste of resources. It is advantageous to improve these properties with antibody engineering strategies. Engineering efforts can be classified into two broad categories: (i) directed evolution and (ii) computational approaches including rational design.

Directed evolution strategies include display systems that allow screening of an antibody library (naïve, immune, synthetic, or semi-synthetic) [94]. Display systems can be classified into two main groups: (i) *in vitro* (phage, ribosome, mRNA display) and (ii) *in vivo* (bacterial, yeast, mammalian surface display) [95]. Phage and yeast display methods have been widely preferred thanks to their easy screening process [96]. A recommendation of the European Union Reference Laboratory for alternatives to animal testing (EURL ECVAM)'s Scientific Advisory Committee (ESAC) states that "when a scientifically valid alternative [to animal-derivation of antibodies] is available, then it simply must be used" [97]. In this context, display systems fulfill and simplify the necessary ethical research guidelines and regulations.

Computational approaches include molecular dynamics (MD) simulations, three-dimensional structure modeling (de novo design) or modeling of antibodies with known sequences without crystallization data (homology modeling) based on existing antibody structures, and antibody-antigen docking [98]. Databases are also needed to design and analyze antibodies. Databases may contain sequence (abYsis), structure (Protein Data Bank, PDB), experimental results (Immune Epitope Database, IEDB), or a combination of these [99-101].

Rational design approaches follow a more targeted route by identifying and improving the specific regions of an antibody structure [95]. With the recent advancements in the field, *in silico* approaches for antibody discovery and development are referred to third generation after first-generation *in vivo* approaches and second-generation *in vitro* approaches [102]. The success of an antibody candidate is defined by the overall favorable combination of its developability properties [103]. Progress in web-based rational design tools can direct the efforts toward problematic regions and help to improve solubility, stability, and aggregation properties (Figure1.6) [95].

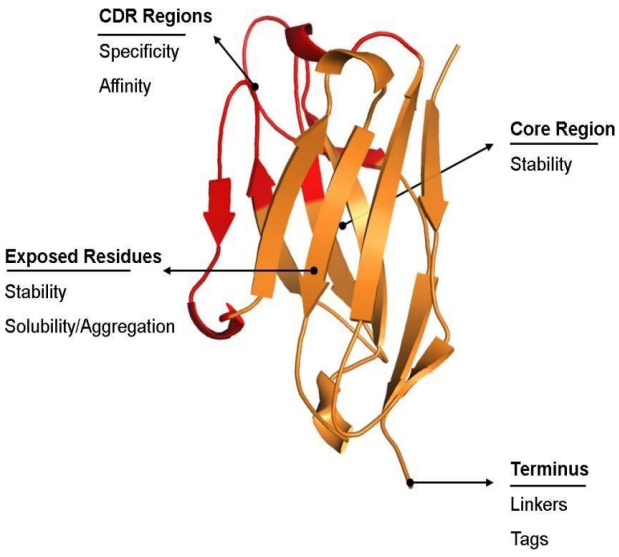


Figure 1. 6 Candidate antibody regions for rational design

Exposed residues can be altered for improved solubility/ aggregation while both exposed and core residues can be altered to increase stability. CDRs primarily impact specificity and affinity, while the inclusion of tags/linkers can be advantageous for enhancing functionality. The figure is adapted from [95].

1.5. Thesis objectives

Targeted therapy is a significant area of focus in the development of biological drugs, and antibodies play a pivotal role as one of the most important agents in this regard. Recent advancements in antibody engineering applications have significantly expanded the repertoire of therapeutic antibodies, with numerous types now progressing through clinical trials or receiving approval. It is important to expand the capability of antibody engineering approaches to provide tightly controlled, reproducible antibody quality defined by developability properties. Antibody developability efforts in the early development phase yield benefits, as these efforts reduce the number of failed antibodies in later stages, therefore preventing the loss of resources and time. Herein, antibody sequence and structure information, and molecular level of understanding of antibody developability are invaluable for the engineering efforts.

Therapeutic strategies through blocking VEGF/VEGFR interaction in cancer are successful that are related to tumor angiogenesis which is a recognized hallmark of cancer. However, the benefits of the single targeting VEGF and the currently approved antibodies are limited. Overexpression of PD-L1 on cancer cells suppresses the anti-tumor response, therefore, blocking PD-L1 becomes important in immune checkpoint blockade strategies. Since VEGF supports the immunosuppressive tumor environment, dual blockade of angiogenesis and immune checkpoints (VEGF and PD-L1) becomes a recent combinatorial cancer therapy with great potential. But a therapeutic antibody that targets these antigens is yet to be discovered. There are a limited number of approved anti-VEGF (Table 2.1, Chapter 2) and anti-PD-L1 (Table 5.1, Chapter 5) antibodies in cancer therapies. Antibody formats, both in constructional and targeting manners, against VEGF and/or PD-L1 remain uninvestigated in terms of the recent antibody engineering applications.

In this thesis, first, I aimed to (i) generate an scFv fragment targeting VEGF (ii) evaluate the role of Vernier zone residues on antibody affinity-stability trade-off using VEGF targeting scFv (iii) improve the developability of the parental anti-VEGF scFv through modulation of non-

CDR regions. Second, I aimed to (iv) enlighten the antibody specificity modulation through dual-specific antibodies, (v) evaluate the role of Vernier zone residues on antibody affinity-specificity trade-off utilizing VEGF and PD-L1 targeting as a case study, (vi) establish a novel antibody engineering strategy to gain dual-specificity through modulation of non-CDR regions, and as a result (vii) generate a dual-specific scFv fragment against VEGF and PD-L1.

1.5.1. Development of anti-VEGF scFv from its full-length antibody

In chapter 2, I present the design of the scFv fragment with two different flexible linkers from full-length humanized antibody Bevacizumab, the methods used to optimize production in bacterial expression, purification, and characterization of scFv to obtain protein in higher yields and purities. Thereafter, I analyzed the effect of the linkers, non-repetitive and repetitive, on scFv characterization. scFv with a non-repetitive linker that developed within the scope of this chapter was chosen as parental scFv for further rational design studies.

1.5.2. Development of anti-VEGF scFv with improved affinity-stability properties

In chapter 3, I present the rationally designed mutations near or within the Vernier zone region of the scFv to improve the antibody affinity-stability profile. I developed several scFv variants from the parental scFv with site-directed mutagenesis, produced in bacterial expression system, purified, and compared the biophysical and biochemical properties with optimized methods within the scope of chapter 2. Next, I present the molecular dynamic studies (in collaboration with Kale Lab, IBG) of the promising scFv variants with improved affinity and/or stability that helps us to enlighten antibody affinity-stability trade-off. Here an improved variant that shows the highest affinity increase to target VEGF, was chosen as parental scFv for further specificity modulation studies.

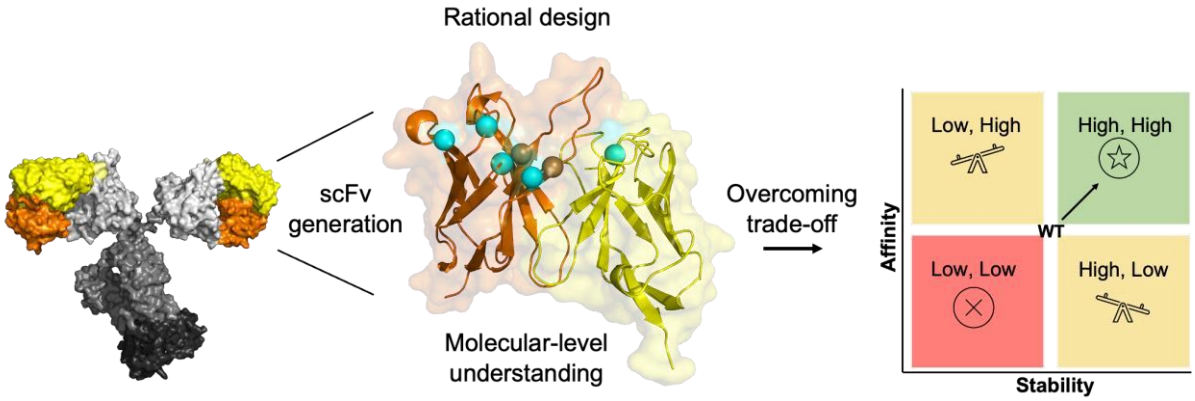


Figure 1. 7 Overview of thesis - 1

1.5.3. A novel specificity engineering strategy: Development of dual-specific scFv via Vernier zone diversification

In Chapter 5, I present a very first proof-of-concept study demonstrating the engineering strategy employed to develop a dual-specific antibody. Some parts of this chapter are carried out in collaboration with Callewaert Lab, VIB-UGent, Belgium. The role of antigen-facing, non-hypervariable loop called light Vernier 4, LV4, evaluated on modulation of antibody specificity through directed evolution approaches. Mono-specific, affinity-stability improved anti-VEGF scFv is used as a template for the studies. Based on the conclusions of Chapter 4 results, a site-saturated synthetic library is designed within the LV4. Library generation methods are optimized to obtain higher transformation efficiencies. The corresponding library is displayed on the yeast surface. Sorting experiments are performed against the second antigen PD-L1. The sequence of the enriched scFv clones was identified. Flow cytometry analyses are carried out to distinguish the inherited and acquired binding profiles of parental and enriched scFvs against antigens. Here an enriched scFv variant that shows significant PD-L1 binding while retaining the VEGF binding is discovered. We show that rationally designed, a small set of mutations within non-CDR regions have great potential to gain new antigen specificity. This novel engineering strategy offers a new road in specificity modulation and can reduce the cost of experimental burden while still requiring follow-up engineering efforts for further improvements.

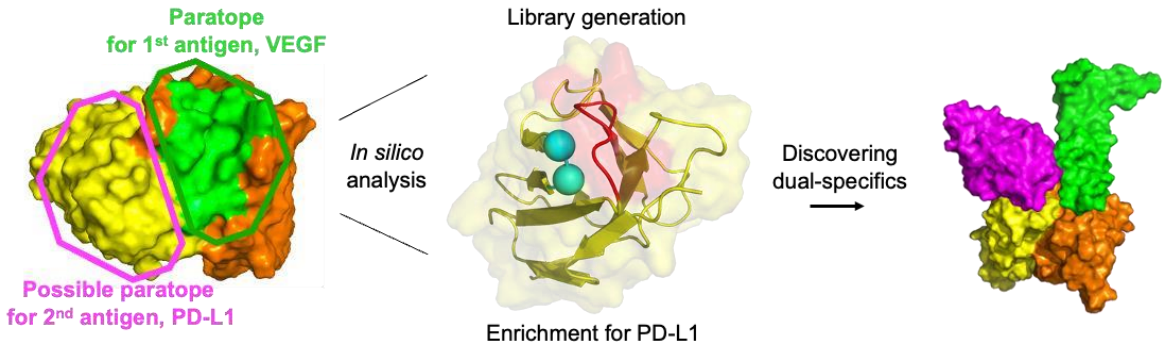


Figure 1. 8 Overview of thesis - 2.

2. Design, production, purification, and characterization of an anti-VEGF scFv

2.1.Introduction

2.1.1. Single chain antibody fragments (scFvs)

Variable fragment (Fv) consists of the variable domains of the heavy and light chains of an antibody which are connected with a disulfide bond [104]. Single chain antibody fragment (scFv) is the engineered version of Fv of an antibody where the variable domains are coupled together via a linker instead of a disulfide bond. ScFv format offers several advantages over a full-length antibody: (i) the small size of around 25 kDa makes them suitable for protein engineering approaches [105], (ii) scFv format does not require PTM that allows them to be suitable for large-scale production in different microbial expression systems, therefore faster, lower cost, in higher yields [106, 107], (iii) since the variable chains linked to each other on gene level, there is no need to balance the expression of variable domains [108], (iv) their small size provides better tissue penetration in cancer immunotherapy compared to larger formats [109].

Omitting the Fc region in the scFv format can be evaluated from two perspectives. Antibody effector functions are induced by the interaction between the Fc region and its complementing proteins or Fc receptors in the immune system. Lack of Fc region makes scFvs advantageous as it still functions as inhibiting/blocking [110, 111] (interrupts its function) or neutralizing [112] (abolishes its downstream cellular effects) antibody without unnecessarily activating immune system [113]. Thus, they are considered less immunogenic [108, 114]. On the other hand, they possess low stability, and high aggregation propensity [115, 116]. Therefore, protein aggregation-mediated immune activation should be taken into account [117]. Also, scFvs have

a shorter half-life (5-6 hours [118]) which requires more frequent dosing [119]. These drawbacks generally can be easily solved via protein engineering approaches such as fusion proteins and rational design [95, 118, 120].

scFvs and its fusion designs have been successfully expressed in several expression systems [121]. *Escherichia coli (E.coli)* is one of the heavily utilized bacterial strains for recombinant protein production. The advantages of using *E.coli* as a bacterial expression can be listed as the species are well-known, readily available, and easy and cheap to handle [122]. It also provides high cell density with fast growth kinetics [123, 124]. scFvs can be produced in the *E.coli* as soluble, functional proteins or aggregated, inclusion bodies. To enable soluble production, an expressed protein is directed to *E.coli*'s periplasmic space through the utilization of a signal peptide [125]. This oxidizing environment includes chaperones and disulfide isomerases [126], which promotes the formation of disulfide bonds which is critical to the assembly of a functional Fv [127]. On the other hand, direct expression of scFvs into the cytoplasm of *E.coli* induces the constitution of insoluble aggregates, called inclusion bodies, requiring *in vitro* refolding for the disulfide bonds formation and correct folding to be functional [128].

The scFv linker is a critical parameter that significantly influences both *in vivo* and *in vitro* properties [129, 130]. The linker between the chains plays a key role for a correct V_H-V_L antigen-binding interface, thereby impacting the overall function of the scFv. Therefore, the design of the peptide linker is crucial for the successful scFv pairing [129, 131]. The overall profile of the scFv, including expression level, affinity/specificity, folding, oligomeric state, stability, and in vivo activities, can potentially be influenced by the length and sequence composition of the linker [131-134]. The Glycine-serine (GS) repeat sequence is widely employed in scFv design, primarily because of its inherent flexibility. Various lengths and combinations of GS linkers have been investigated for scFv fragments, with $(G_3S)_n$ and $(G_4S)_n$ motifs being the most prevalent one's [130, 135-137]. The optimal length of the linker can be tailored within the range of 5 to 35 amino acids to enhance the performance of scFvs for diverse applications [131, 138]. If the scFv linker exceeds 12 residues in length, the genetically fused V_H and V_L domains form an operating scFv, typically exhibiting high monomeric properties. On the other hand, scFvs with shorter linkers, shorter than 12 a.a., might form multimers by interacting with other scFv molecules [139, 140]. Although GS linkers offer the desired flexibility for correct scFv folding, their repetitive nature can pose challenges in PCR-based

engineering strategies [141] and may induce immunogenic responses [142]. Hence, alternating non-repetitive linkers that offer similar flexibility, might be employed to enhance *in vitro* properties of scFvs.

2.1.2. Angiogenesis and Anti-Vascular Endothelial Growth Factor (VEGF) strategies

Angiogenesis is a complex process of forming new blood vessels from existing ones, occurring in multiple stages. New blood vessel formation is important for normal developmental stages [143], wound healing [144], and the female reproductive cycle in adulthood [145]. However, it is deregulated in various diseases [146-149]. Also, excessive, or abnormal angiogenesis contributes to the tumor growth which is established as a hallmark of the cancer [150]. The high level of oxygen and nutrient needs in tumor cells cannot be met by the existing vascular network. This initiates the secretion of angiogenic growth factors [151]. The most found tumor-secreted angiogenic molecule is VEGF. VEGF is upregulated in tumor cells and binds its cognate tyrosine kinase receptors (VEGFRs) on endothelial cells. This leads to downstream signaling for endothelial cell proliferation and migration, matrix reorganization, and degradation, therefore tumor angiogenesis, metastasis, and growth [152, 153]. VEGF is a member of a family that includes subtypes of VEGF-A, VEGF-B, VEGF-C, VEGF-D, and placental growth factor (PIGF). VEGF-A is the most prominent member of the family. Due to its dominant role in VEGF-related pathways, VEGF-A is usually named VEGF in the literature [154].

VEGF is a critical molecule in cancer angiogenesis. Thus, blocking the binding of VEGF to its receptor suppresses the tumor growth [155]. As a result, there has been significant elevation in the development of therapeutic approaches targeting VEGF, and numerous drugs targeting VEGF are currently in use. [156, 157]. Currently there are 5 approved antibodies (out of 180) and 8 antibodies (out of 138) in late-stage clinical studies that blocks the VEGF/VEGFR interaction for ophthalmologic and cancer therapies (Table 2.1 – Accessed date 21 April 2023) [158]. Bevacizumab (Avastin, Genentech) is the first approved anti-VEGF monoclonal antibody in 2004 [159].

Bevacizumab is first reported as the humanized mouse anti-VEGF antibody by site directed mutagenesis of several framework and CDR residues. Bevacizumab binds to all human VEGF isoforms (VEGF-A) with 0.5 nM affinity [17]. The bevacizumab binding epitope of VEGF was

determined by x-ray crystallization as a Fab-ligand complex [160], that overlaps with the receptor binding epitope of VEGF (VEGFR2) [161]. It has been shown that Bevacizumab suppresses the tumor growth in preclinical models by regressing and normalizing of existing vessels, inhibiting new vessel growth [17, 162, 163]. However, the best efficacy results of Bevacizumab have been reported when it is combined with chemotherapy agents [164-166].

Table 2. 1 VEGF/VEGFR blocking antibodies approved and in late-stage clinical studies

Approved antibodies				
Antibody	Target	Format	Specificity	Therapeutic Area
Bevacizumab	VEGF	Full-length	Monospecific	Cancer
Brolucizumab	VEGF	scFv	Monospecific	Ophthalmology
Faricimab	VEGF	Full-length	Bispecific (Ang-2)	Ophthalmology
Ramucirumab	VEGFR	Full-length	Monospecific	Cancer
Ranibizumab	VEGF	Fab	Monospecific	Ophthalmology
Antibodies in late-stage clinical studies				
Antibody	Target	Format	Specificity	Therapeutic Area
Sevacizumab	VEGF	Full-length	Monospecific	Cancer
JY-025	VEGFR	Full-length	Monospecific	Cancer
601	VEGF	Full-length	Monospecific	Ophthalmology
9MW0211	VEGF	Full-length	Monospecific	Ophthalmology
ABL001	VEGF	Full-length	Bispecific (DLL4)	Cancer
Ivonescimab	VEGF	Full-length	Bispecific (PD-1)	Cancer
Navicixizumab	VEGF	Full-length	Bispecific (DLL4)	Cancer
Tarcocimab	VEGF	Full-length	Monospecific	Ophthalmology

Since the approval of the first anti-VEGF antibody Bevacizumab, induced local invasiveness, distant metastasis is related to anti-angiogenesis approaches [167, 168], tumor cells evolved to resistance to anti-angiogenesis therapies [169]. VEGF role is established as an immunosuppressive in tumor microenvironment [170-172]. Also, powerful protein engineering approaches applied for antibodies, different constructional formats (such as Fab, scFv-Fc) and targeting formats (such as diabodies, bispecific) of antibodies are developed to a greater extent. Thus, the number and benefits of the current anti-VEGF antibodies remained limited and uninvestigated in terms of the recent developments.

2.1.3. Chapter overview and publications

Due to their smaller size and the ability to apply diverse engineering techniques, scFvs are highly preferred in both diagnostic and therapeutic fields. Despite sharing the same sequences of the variable chains, biophysical and biochemical properties of scFv generally do not match to parent antibody due to lack of the constant domains and structural smallness. Also genetically linking of the variable chains can alter the domain dynamics and relative orientation.

The linker between the variable chains in scFv plays a crucial role by covalently joining these domains, and it can significantly impact the biophysical, chemical properties, and *in vivo* activity of the scFv. Consequently, the design of the scFv linker is a vital aspect in ensuring successful scFv construction, with the preference for flexible linkers that facilitate the correct pairing of variable chains. The flexibility of the linker is defined by its length and sequence composition. Here, glycine-serine (GS) linkers are favored in scFv design due to their high flexibility. However, despite the flexibility advantage provided by GS linkers, their repetitive sequence poses challenges for PCR-based engineering strategies and may also contribute to immunogenicity.

In this chapter, we have developed scFv constructs using two distinct types of linkers: a repetitive GS linker and an alternative non-repetitive linker. We designed scFv fragments from commercial anti-VEGF humanized antibody, Bevacizumab. We have optimized (i) the production of the scFv in *E.coli* expression system, (ii) the purification via affinity chromatography (his-tag and Protein L), (iii) biophysical and biochemical characterization. Finally, we have compared the effect of two different linkers on scFv characteristics utilizing optimized methods within this chapter. The results from this chapter highlights that the non-

repetitive linker showed a better *in vitro* profile such as a higher monomer ratio, higher thermal stability, and lower immunogenicity.

The results from this chapter were published in Scientific Reports as an original research paper in 2022 [173]. This study was also selected for poster presentation in BIO-Turkey International Biotechnology Congress 2021, 9-11 Sept 2021.

2.2.Material and Methods

2.2.1. scFv construction

scFv fragment of the full-length anti-VEGF antibody was designed from bevacizumab (Brand name Avastin), and the sequence was obtained from the database, DrugBank with an accession number DB00112 [174]. Antibody residues were numerated according to the Kabat numbering. Variable domain residues were determined by PyMol analysis of the crystal structure of Bevacizumab, 1BJ1 [160]. Seventeen amino acids length of the linker was chosen from a previously described non-repetitive flexible linker library with a one amino acid difference (SPNSASHSGSAPQTSSAPGSQ) [141]. Variable domains fused with this linker, were cloned into pET17-b expression plasmid with a leader sequence (PelB), FLAG-tag, and hexa-histidine-tag. The same scFv construct is also generated with a repetitive GS linker for further comparisons, named scFv-L2. These scFv constructs in pET-17b expression vectors are ordered from GenScript.

2.2.2. Cell growth, carbon utilization, and protein expression

Heat shock protocol (42 °C for 42 seconds) was performed for scFv plasmid transformation into *E.coli* strain BL21 (DE3) PLysS (Thermo Scientific, EC0114). It was plated on LB agar, Fisher plate containing 100µg/mL ampicillin (plasmid resistance) and 25µg/mL chloramphenicol (cell resistance). These antibiotic contents were used for all growth steps. Agar plate was incubated at 37 °C for overnight (16-18 hours). A single colony was picked from agar plate for expression, plate was stored in the 4°C refrigerator up to one week. Picked single colony was inoculated into 5 mL LB broth and incubated at 37 °C for overnight (16-18 hours) with shaking at 250 RPM. 1 mL of this starter culture was mixed with 50% sterile glycerol, stored at -80°C freezer for further growths.

Absorbance at 600 nm of this starter culture was measured, required volume of the starter culture to start growth at 0.05 of OD_{600} in 300 mL growth medium was calculated. Calculated volume was inoculated into 300 mL auto-induction medium [175] (1 % tryptone, 0.5 % yeast extract, 25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM Na_2HPO_4, 1 mM MgSO4, 0.5 % glycerol, 0.05 % glucose, 0.2 % α-lactose) in a one-liter flask, grown at 10 ℃, 18 ℃, 30 ℃ for 48 hours with shaking at 250 rpm. After 48 hours growth, cells were pelleted at 10000 x g for 20 minutes at 10 ℃ (Avanti), then the supernatant was collected for purification, stored at 4 ℃ up to one week.

Carbon utilization determination by HPLC

Time dependent change in the amount of carbon sources in auto-induction medium (glucose, glycerol, lactose, galactose, acetic acid, ethanol) were monitored via an organic acid column (Transgenomic ICSep ICE-ION-300, ICE-99-9850) in HPLC (Shimadzu LC-2010) with refractive index detector (RID-10A). Standards were prepared in a range of 0.15-20 g/L (0 - 2000 ppm) except glycerol (0 – 5000 ppm). Retention time and peak area count of standards were determined in 0.0085 N H_2SO_4 at oven temperature 70 ºC, flow rate 0.4 mL/min for 40 mins. Supernatant samples were collected every 6 hours, filtered through 0.45 µm before analysis. The amount of the carbon sources at each datapoint were calculated according to the standard curves that were determined based on the peak area count at increasing concentrations.

2.2.3. Protein purification and detection

pH of supernatant was arranged to 7-7.4 with 5 M NaOH solution for histidine tag purification steps. pH adjusted supernatant was either loaded to prepacked nickel sepharose affinity resin column, HisTrap, (Cytiva, 17524802) or incubated with 1 mL of His-Pur Ni-NTA resin (Thermo Scientific, 88221) per 100 mL supernatant. Supernatant was applied directly onto the column, washed, and eluted via peristaltic pump with a 2.5ml/min flow rate. Resin-supernatant mixture was incubated for 2 hours or overnight and loaded onto a 10 mL vacuum column. Purification steps proceeded according to manufacturer protocol. 20 mM KH_2PO_4, 500 mM NaCl, 25- or 40-mM imidazole, pH 7.4 buffer was used as binding and wash buffer, 20 mM KH_2PO_4, 500 mM NaCl, 250-500 mM imidazole, pH 7.4 buffer was used as elution buffer.

Eluted protein sample was buffer exchanged into PBS, (1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 2.7mM KCl, 137 mM NaCl) and concentrated through membrane filtration (Amicon® Ultra-4 Centrifugal Filter Units, MWCO 10 kDa, Merck, UFC8010). The protein amount was

determined using a calculated extinction coefficient and molecular weight of scFv by Nanodrop 2000 at 280 nm.

After collecting the histag-purified protein samples from 3-5 expression, either unfold-refold method or protein L affinity chromatography was performed to achieve better purity (>95%). For unfolding-refolding; the histag-purified protein sample volume was arranged to 100 mL with 1X PBS, then arranged to 200 mL with binding buffer containing 20 mM imidazole and 6M GdnHCl. 200 mL unfolded-sample was incubated with 1 mL Ni-NTA resin for 2 hours at +4 ºC, loaded onto a 10 mL vacuum column. Protein was eluted with 10 mL elution buffer containing 250 mM imidazole and 6 M GdnHCl. Elution sample was loaded into MWCO 3.5 kDa dialysis membrane (Serva, 4431001), membrane left into 3-liter 1X PBS at +4 ºC for overnight, concentrated through MWCO 10 kDa membrane filtration. For Protein L binding, histag-purified protein sample was diluted with binding buffer to 10 column volume (CV), applied directly onto the prepacked Protein L column (Cytiva, 29048665), washed, and eluted via peristaltic pump with a 1 ml/min flow rate. Purification steps proceeded according to manufacturer protocol. 100 mM Na_2HPO_4, 150 mM NaCl, pH 7.2 buffer was used as wash buffer, 100 mM glycine pH 2.6-2.7 was used as elution buffer, 1 M Tris-HCl at pH 9 was used as neutralization buffer. Eluted protein sample buffer exchanged into PBS and concentrated through MWCO 10 kDa membrane filtration. The protein amount is determined using a calculated extinction coefficient and molecular weight of scFv by Nanodrop 2000 at 280 nm.

SDS-PAGE and Coomassie staining was used to assess the purity of the protein. SDS gel was prepared according to TGX™ FastCast™ Acrylamide Kit, 12% protocol (Bio-rad, 1610175).

Coomassie staining for 1 hour and de-stained with dH_2O for overnight. Gel was visualized in Bio-rad. Western blot analysis was used to detect the target protein scFv in purified protein sample. Mini-PROTEAN® Tetra Cell and Mini Trans-Blot® Module system (Bio-Rad, 1658029) was used to perform Western Blot. 1/10000 diluted Mouse anti-flag and 1/10000 diluted anti-mouse hrp conjugated antibodies were used primary and secondary respectively. ECL substrate (Thermo Scientific, 32106) and 1M H_2SO_4 were used for visualization in Bio-Rad ChemiDoc Imaging System.

2.2.4. Biophysical and biochemical characterization

Monomer/dimer ratio determination by HPLC-SEC

Monomer/dimer ratio of the soluble scFv in PBS was determined via an size exclusion column (Tosoh Bioscience, TSK-gel SuperSW3000 column, 18675) in HPLC (Shimadzu LC-2010) with UV-VIS detector. Retention time of standards (thyroglobulin, immunoglobulin G, bovine serum albumin, myoglobin, uracil) were determined in phosphate buffer (100 mM KH_2PO_4, 100 mM Na_2HPO_4, 100 mM K_2SO_4) at oven temperature 25 ºC, flow rate 0.3 mL/min for 20 mins. The injected amount of scFv was 5 µg at the same conditions with standards. The absorbance values were monitored at 280 nm. Retention time of scFvs was determined using the molecular weight of standards.

Thermal melting point determination by Differential Scanning Fluorimetry (DSF)

Thermal shift studies of pure scFvs were determined with a hydrophobic dye, SYPRO Orange dye (Sigma, S5692) via ABI 7500 Fast RT-PCR, in continuous ramp mode at 0.05% ramp rate. Fluorescence signals were recorded in the melt curve filters, 580 ± 10 nm for the excitation filter and 623 ± 14 nm for the emission filter, between 25-99°C. For optimal, concentrations of dye and protein were determined as 2x and 2 □M, respectively. Thermal melting points (T_m values) from the Hill equation fitted data were defined via the in-house python scripts.

Binding kinetic determination by SPR and ELISA for affinity

Surface Plasmone Resonance (SPR)

Binding affinity kinetics were determined using a Biacore T200 instrument. VEGF at 1000 nM concentration was immobilized on a CM4 chip (Cytiva, 29104989), at a flow rate of 10 µl/min for around 1 min in HBS-EP (0.01 M HEPES, 0.15 M NaCl, 0.0003 M EDTA and 0.005% v/v Surfactant P20) buffer, pH 7.4 (Cytiva, BR100188). Target RU was determined at 300 RU for VEGF immobilization. In a range of 0-100 nM scFv were subsequently injected at a flow rate of 30 µl/min onto the VEGF-coated surface in HBS-EP buffer, pH 7.4 with 0.05% w/v BSA. The data obtained were corrected by comparing them to a control flow cell that did not contain VEGF. Additionally, the data were further corrected by comparing them to the flow cell with buffer injection. Data from the SPR curves were analyzed using the manufacturer's software, BiaEval 3.0. K_D, k_{on} and k_{off} values were calculated by fitting curves to a 1:1 binding model.

SPR curves were extracted from the software and plotted using in-house python scripts.

Enzyme-Linked ImmunoSorbent Assay (ELISA)

Apparent biding affinity was determined using the ELISA assay. Each well in high protein binding 96-well plate was immobilized with 1 ng VEGF (Gibco, PHC9394) in 100 uL of PBS, at +4 ºC for overnight. On the morning of the following day, wells were incubated with 100 µl of 10% w/v skim milk PBS blocking solution with 80 rpm agitation (orbital shaker) at room temperature for 6-8 hours. Wells were washed, in a range of 0 nM to 1000 nM scFvs in 100 uL PBS and were added into wells as three replicates at +4 ºC for overnight. On the morning of the following day, wells were washed, and subsequently incubated with 1/5000-10000 diluted mouse anti-flag IgG antibody and 1/5000-10000 diluted anti-mouse IgG HRP conjugated antibody. Afterward, wells were incubated with TMB substrate solution (Thermo Scientific, 34028) at room temperature for 15-30 minutes depending on color change in the wells. Next, 100 µl of 1M H_2SO_4 as a stop solution was added to wells and the plate was read at 450 nm by a plate reader. Absorbance units were background (no antigen) corrected. Antibody incubation occurred in 100 µl 1% w/v skim milk PBS-T solution with 80 rpm agitation at room temperature for 1 hour, wells were washed between antibody incubations. Three times 300 µl 1X PBS washing is used in all washing steps. The K_D value from the ELISA absorbance readings was determined using the Hill1 equation fit by utilizing the in-house Python scripts.

Aggregation analysis of scFv

scFvs (0.5 mg/mL, 25 µL, in 1X PBS) were incubated at 60 °C, and 220 rpm agitation on a heat block. At several time points (0-420 minutes), protein aliquots were centrifuged at 17,000 g at 4°C for 10 minutes. Protein concentrations of soluble fractions were measured using NanoDrop 2000 (absorbance at 280 nm). The aggregation kinetic value from the absorbance readings was calculated fitting the data to single exponential function (y = a(1-exp(-bx)), where a is the final amplitude, b is apparent aggregation rate constant ($_{kapp}$, s^{-1}) and x is time [176]. Percent insoluble aggregation was calculated by subtracting soluble protein concentration from total concentration after thermal (60°C) and mechanical (220 rpm) stress for 4 hours [173].

2.2.5. *In silico* analysis

The linker's flexibility scores were derived according to the average flexibility index of amino acids [177, 178]. Flexibility values were determined according to amino acid content and length

of the Linker 1 (L1) and Linker 2 (L2). The linker's immunogenicity scores were predicted via the webserver, IEDB T Cell Epitopes Immunogenicity Prediction Tool (http://tools.immuneepitope.org/main/) [179].

2.3.Results

2.3.1. Analysis of Bevacizumab variable domains

Bevacizumab is a full-length IgG antibody, developed to neutralize VEGF. Complex structure of Fab fragment with the ligand VEGF is determined and published in Protein Data Bank with the PDB ID 1BJ1. 1BJ1 contains homo-dimer VEGF molecules, each VEGF pair with the Fab fragment of the Bevacizumab. Variable and constant regions are linked to each other with a short linker region (8 and 7 residues for heavy and light chain, respectively). The C-terminal of the sequences of the variable domains are determined after the last β-sheet and the first 2 residues of the linker region for heavy and light chains (Figure2.1A). The full sequence of variable domains is extracted from DrugBank databank with the accession number DB00112. Variable domain sequences are shown in Figure 2.1B and 2.1C for light and heavy chains. Residues are numarated based on the universal Kabat numbering system [180].

2.3.2. scFv design and construction of expression vector

Light and heavy chains pair with covalent disulfide bonds and non-covalent interactions. Variable domain pairing is driven by only non-covalent interactions. Generally, a flexible peptide linker is used to covalently link to variable domains of the heavy and light chains to form proper scFv pairing. For this scFv design, seventeen amino acids length of linker is chosen from previously described non-repetitive flexible linker library [141]. pelB is a 22 amino acids long peptide sequence which, when genetically fused to a protein, directs the protein to the bacterial periplasm and cleaved by the organism before the secretion of the target protein [181]. pelB sequence is fused to N-terminal of the scFv to increase secretion to supernatant. Commonly used Flag-tag and hexa-histidine tag are fused to C-terminal of the scFv for further purification, detection, and characterization steps. Factor Xa protease site is added between the C-terminal of the scFv and the tags to dispose of these tags in any case requiring scFv protein without tags. Designed scFv is cloned into multiple cloning site of common pET-17b expression vector via NdeI and XhoI restriction sites to use in bacterial expression system (GenScript). pET17-b expression vector contains T7 promoter which allows lac-operon induction, ampicillin

resistance gene for selection and and N-terminally T7 tag. Designed scFv expression vector map is shown in Figure 2.2.

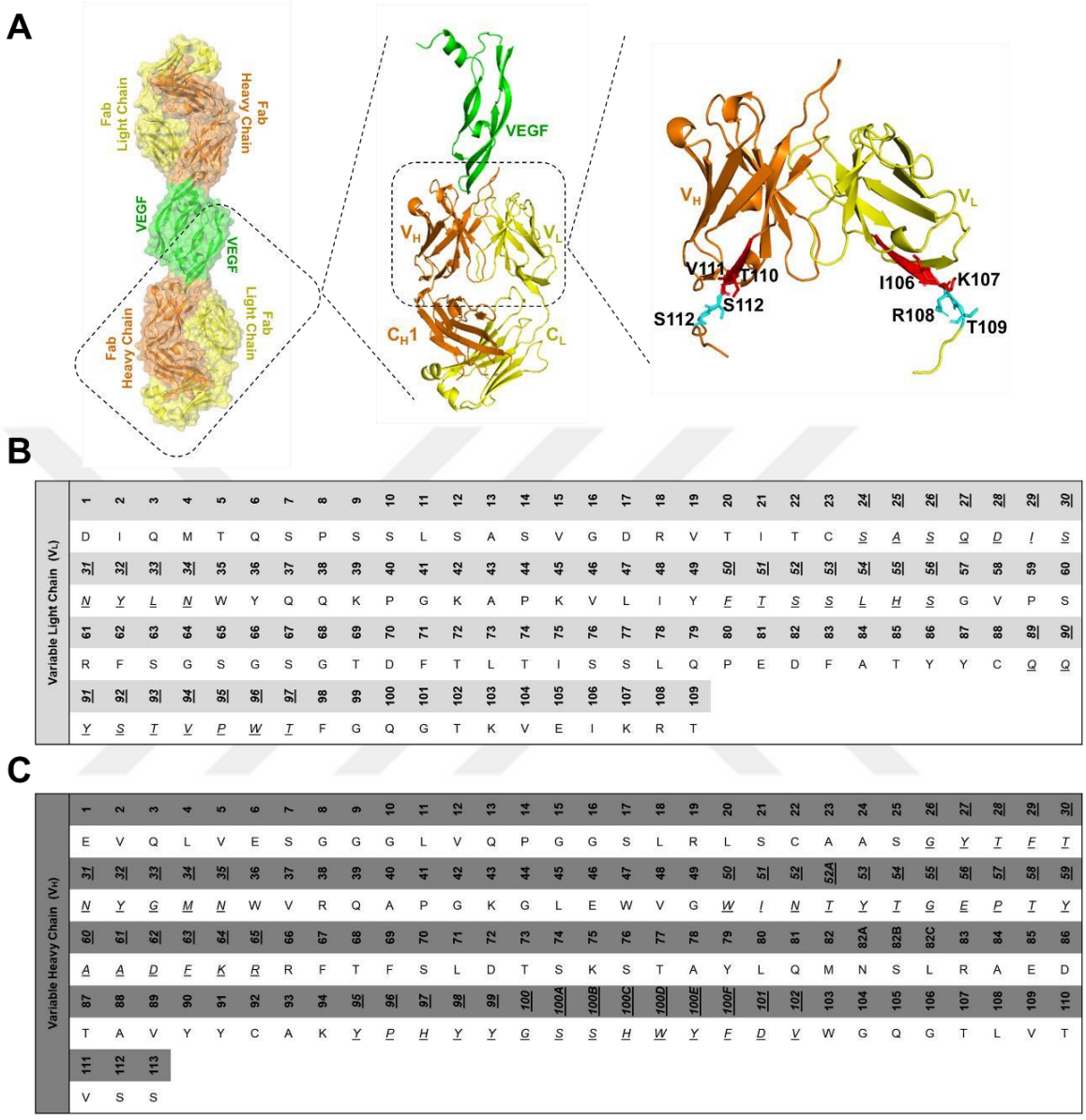


Figure 2. 1 Variable domains of Bevacizumab

A) View of the antigen VEGF bound to Fab fragment of Bevazicumab Inset shows the scFv fragment of the crystal structure shown in cartoon. The C-terminal amino acids of the variable chains are indicated in red, the N-terminal amino acids of the linker region between variable and constant regions are indicated in cyan. B-C) Variable chain sequences of the Bevacizumab. Variable light chain (B) and variable heavy chain (C) are numbered according to Kabat numbering system, CDRs are indicated italic and underlined.

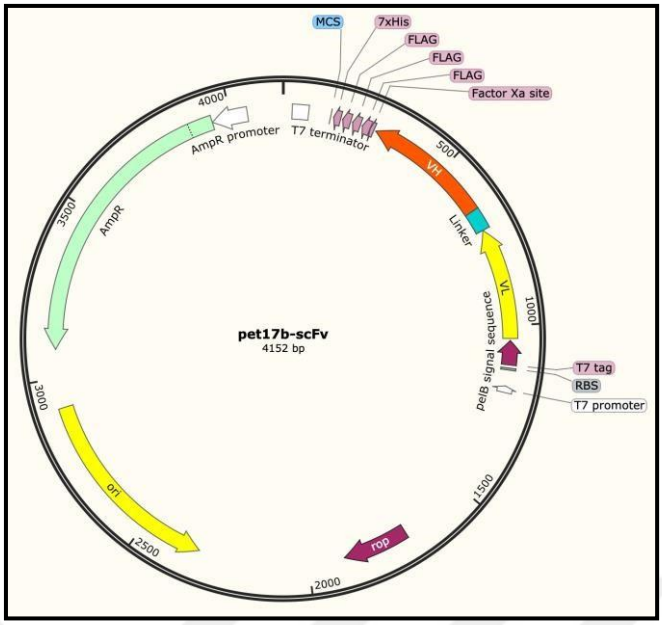


Figure 2. 2 scFv expression vector

Protein parameters are determined based on the sequence of secreted scFv into supernatant which does not include the PelB sequence via Expasy-ProtParam web-tool. Parameters are listed in Table 2.2.

Table 2. 2 scFv protein parameters

Number of amino acids	296	Molecular weight (Da)	32635.56
Number of atoms	4442	Extinction coefficient*	68550
Formula	$C_{1441}H_{2136}N_{390}O_{468}S_7$	Theoretical pI	5.60

*Extinction coefficient (EC) is in units of $M^{-1}cm^{-1}$, at 280 nm measured in water.

2.3.3. scFv expression and purification

Three temperatures 10 °C, 18 °C, 30 °C and two pH ranges 6.8 and 7.8 were screened in autoinduction medium. Highest protein secretion was detected in the supernatant at 18 °C, pH 6.8 and 30 °C, pH 7.8 while higher amount of cell proteins was detected at 30 °C, pH 7.8. There were no growth and protein expression at 10 °C in both pH ranges. Although 30 °C, pH 7.8.

yielded similar to 18 °C, pH 6.8, 18 °C, pH 6.8 were selected as optimized scFv expression conditions considering higher temperatures accelerate the protein degradation (Figure 2.3).

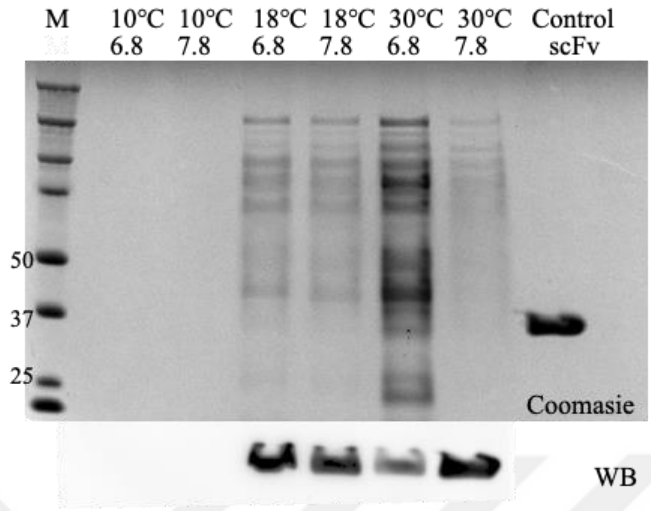


Figure 2. 3 scFv expressions in different temperature and pH conditions

Expression level of the scFv was compared at three temperatures and two pH ranges. Purified scFv molecular weight 32 kDa. Western blot is labelled as WB.

Time dependent change in the amount of carbon sources in auto-induction medium were monitored via refractive index detector in HPLC. The retention time and standard curve of the carbon sources were summarized in Table 2.3. The time-dependent change in the amount of carbon sources in supernatant was calculated using the determined standard curve equation of each compound.

Glucose in the supernatant was consumed from the beginning of the growth and run out after 12-18 hours. Lactose consumption started after or parallel to glucose consumption with the absence of glucose and run out between the 24-30 hours. As hypothesized, galactose and glucose increased as byproducts of lactose, then used as carbon source and run out between the 36-48 hours. Glycerol consumption started after 20 hours of the growth. The level of the all the carbon sources were monitored near-zero at the 48 hours of the growth. **(**Figure 2.4A). The scFv expression was observed parallel to growth in western blot analysis (Figure 2.4B). The induction of the protein expression started after the 24 hours of growth and the intensity of the protein level in the supernatant increased in the process of the time. However, the intensities of

the protein level after 42 hours of growth were found similar. Due to the lack of lactose, and fixed intensity of the protein level in the supernatant, 48 hours of growth was determined for the termination of the growth.

Table 2. 3 HPLC standard curves

Compound	Retention time (min)	Standard curve *	R^2
Lactose	13.46	y = 429.30x + 2272.02	0.99
Glucose	15.94	y = 404.68 + 10427.00	1
Galactose	17.17	y = 416.88x + 7574.70	0.99
Glycerol	22.76	y = 332.72 – 4556.10	0.99

* x : concentration (ppm) y: peak area

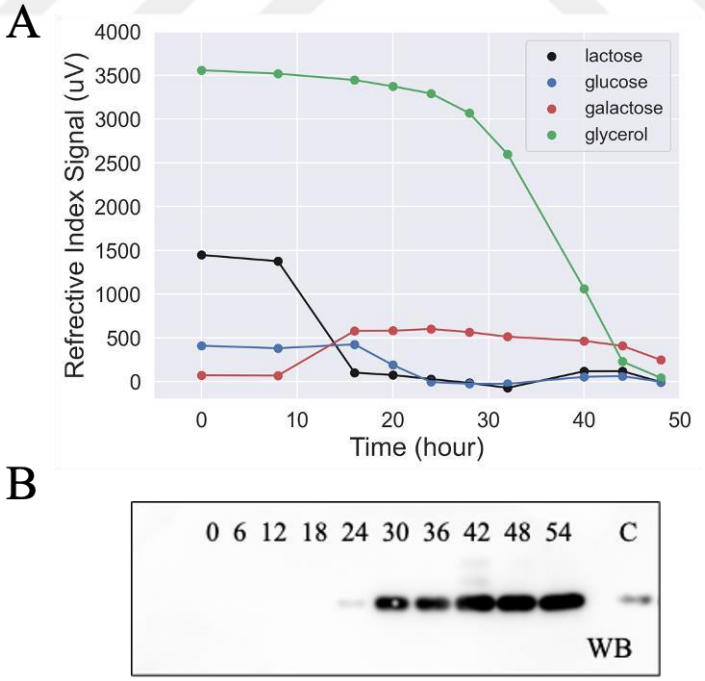


Figure 2. 4 HPLC analysis of consumed carbon sources

A) Time-dependent change of carbon sources in auto-induction medium. B) Time-dependent change (0h-54h) of expressed scFv protein in supernatant, purified scFv was used as a control, labelled as C in western blot (WB).

Hexa-histidine tag (his-tag) of scFv protein was utilized for the first step of purification. Two approaches, Ni+2 affinity Sepharose column (HisTrap) and His-Pur Ni-NTA resin, were used interchangeably. Both approaches yielded similar (4.2 mg protein/ liter culture and 3.7 mg/ liter culture for column purification and resin purification, respectively) and showed same impurity profile in SDS Page (Figure 2.5A). To decrease the impurities in His-tag purification steps, two approaches were utilized, (i) increasing imidazole (ImH) concentration of binding buffer to decrease the non-specific bindings, (ii) 10% glycerol addition to binding (containing 25 mM imidazole) and elution buffer (containing 250 mM imidazole) that glycerol addition might decrease the interaction between scFv and impurities and the non-specific interaction between resin and impurities. Increasing imidazole concentration in binding buffer resulted lower concentrations of eluted fraction (Figure 2.5B) while glycerol addition did not decrease the impurity level in purified sample (data not shown). Binding buffer containing only 25 mM imidazole was used for all his-tag purifications.

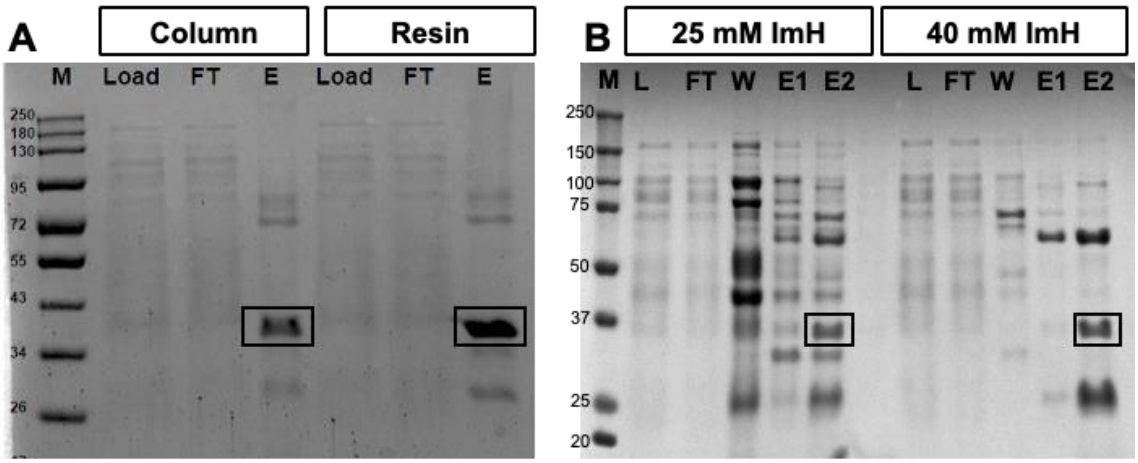


Figure 2. 5 SDS-PAGE profiles of the his-tag purification

A) Comparison of the prepacked his-tag column and his-tag resin purifications. B) Comparison of the different ImH concentrations on purity profile.

To increase the purity of scFv obtained from his-tag purification, two approaches were considered, (i) unfold-refold of proteins, (ii) protein L purification. Both approaches were increased the scFv purity over 95%. However, unfold-refold method resulted degraded scFv and decreased the long-term stability of the scFv (Figure2.6A). Protein L affinity chromatography yielded higher purity, no detectable degradation and scFv recovery calculated over 65% (Figure2.6B). In this experiment, we successfully expressed and purified scFvs with both non-repetitive and repetitive linkers.

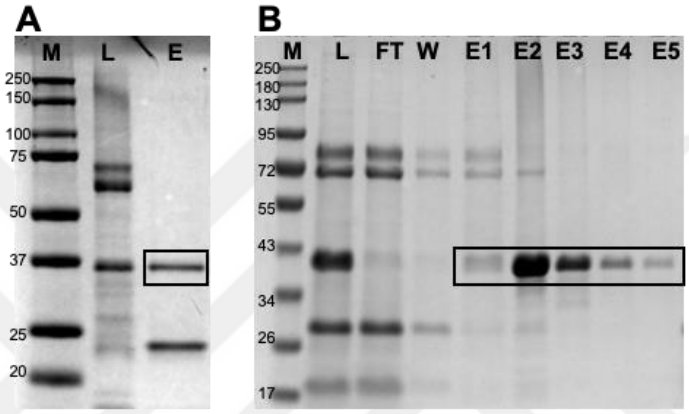


Figure 2. 6 Polishing approaches to increase protein purity

A) Unfold-refold method. B) Protein L purification. M: Marker, L: Load, FT: Flow-through, E, E1-5: Elution fractions.

2.3.4. scFv Characterization

Multimeric formation of the scFv was determined using HPLC-SEC. The dimeric and monomeric forms were distinguished based on their respective retention times, which were determined to be 11.79 ± 0.05 and 13.34 ± 0.06, respectively. The distribution of multimeric and monomeric forms in the solution was quantified by analyzing the peak area count. The results revealed that the monomeric form accounted for 95.69 ± 0.70%, while the dimeric form constituted 4.31 ± 0.64% of the total (Figure 2.7A). The thermal stability of scFv was determined by diffential scanning fluorimetry. Thermal melting point of the scFv was determined at 51. ± 0.4 °C. (Figure 2.7B). Aggregation propensity of scFv was monitored under heat and mechanical stress conditions. By assuming that the protein loss corresponds to the aggregated fraction, the aggregation kinetic value was determined as 0.0011 s^{-1} based on the

absorbance readings (Figure 2.7C). Binding kinetics of scFv was analyzed based on their bindings to their ligand, VEGF. SPR was used and corresponding association (k_{on}), dissociation (k_{off}) constants, and binding affinity (K_D) were determined $3.9x10^{4}$ ($M^{-1}s^{-1}$), $9.9x10^{-5}$ (s^{-1}), 2.51 $x10^{-9}$ (M) (Figure 2.7D).

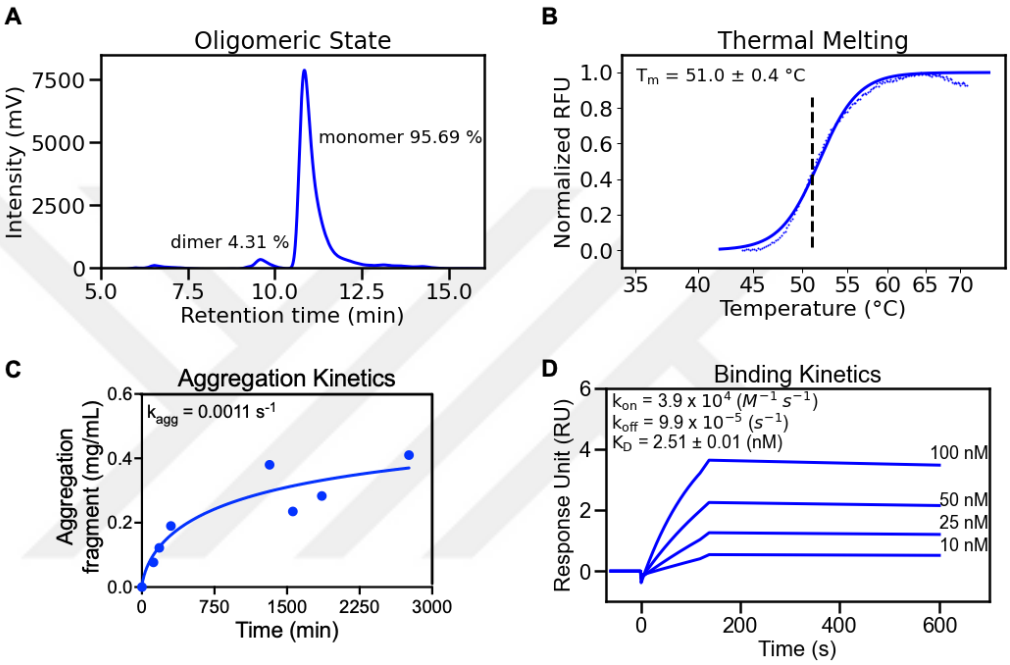


Figure 2. 7 Biophysical and biochemical characterization of scFv

A) SEC chromatogram of scFv from SE-HPLC analysi B) Thermal melting point of scFv from DSF C) Aggregation kinetics D) Binding kinetics.

Furthermore, scFv with non-repetitive linker named scFv-L1 compared to scFv with repetitive GS linker named scFv-L2 in terms of their biophysical and biochemical characteristics, results are summarized in Table 2.4. Both linkers have sufficient amino acid lengths for correct structural pairing of variable chains to form monomeric scFv. Flexibility of the linkers were assessed via calculating the average of the flexibility score of each amino acids within the linker sequences, and both linkers showed close flexibilities. Immunogenicity scores were determined

by IEDB tool [179]. A higher score indicates the bigger probability of starting an immune response. scFv-L2 showed a higher immunogenicity score than scFv-L1.

2.3.5. Linker comparison

scFv-L1 and scFv-L2 were expressed, purified with high purity and characterized as described above. The distinction in molecular weight, around 1 kDa, resulting from the difference in linker sequences, was observed and highlighted through Coomassie staining (Figure 2.8A). The linker effect on the multimeric formation of the scFvs was determined calculating the peak area count by HPLC-SEC. Significant monomer percentage difference was observed between scFvs, where scFv-L1 showed more than 95% monomer, scFv-L2 showed around 66.5% monomer (Figure 2.8B). However, their aggregation profiles were similar (Table 2.4). Thermal melting points were determined 51.4 °C and 50.2 °C for scFv-L1 and scFv-L2, respectively. This showed that scFv-L1 was thermally more stable than the scFv-L2 (Table 2.4C). Binding kinetics of were determined as 0.38 nM, and 2.51 nM, for scFv-L1, scFv-L2, respectively (Table 2.4C). scFv-L2 showed slightly better binding kinetics than the scFv-L1.

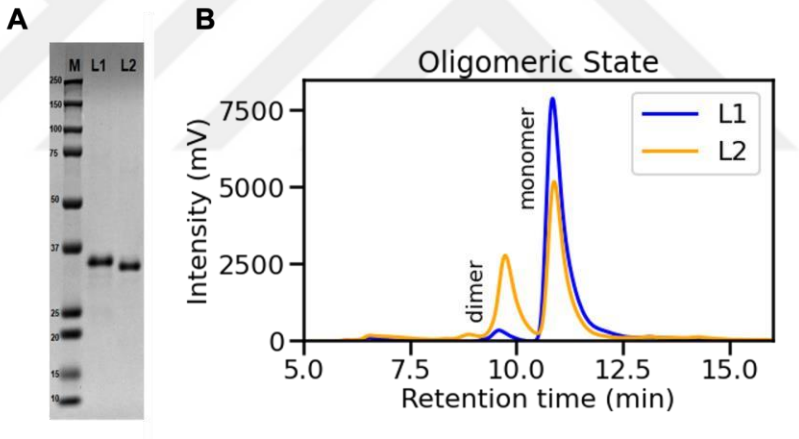
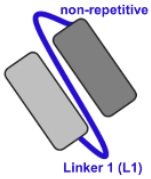


Figure 2. 8 In vitro profiles of the scFvs

A) SDS-PAGE profiles of the scFvs, B) Oligomeric states of the scFvs

Table 2. 4 Properties of scFv-L1 and scFv-L2

Characterization	scFv-L1	scFv-L2

Schematic representation		
Linker sequence	SPNSASHSGSAPQTSSAPGSQ	GGGSGGGSGGGSGGGS
Molecular Weight (kDa)	32.8	31.9
Flexibility score	0.47	0.53
Immunogenicity score	-1.09	-0.06
Monomer %	95.7 ± 0.7	66.5 ± 3.2
Insoluble Aggregates %	88.5 ± 2.8	90.2 ± 4.3
Thermal Stability (°C)	51. ± 0.4	50.2 ± 0.6
Binding Affinity (nM)	2.51 ± 0.01	0.83 ± 0.02

2.4. Discussion

ScFvs offer several advantages when compared to their full-length counterparts. These include cost-effective and large-scale production using bacteria, as well as the availability of diverse *in vitro* display technologies that facilitate the enhancement of various antibody characteristics. Consequently, the number of scFvs in both clinical trials and the market is in rapid increase, driven by these advantageous features. Biophysical and biochemical properties of scFvs are mainly determined by their sequence and structure. The linker between variable domains of scFv genetically connects them and has the potential to affect scFv properties. Thus, linker design is an important parameter for the scFv characteristics. GS linkers are the most preferred linkers because of their demonstrated flexibility. However, the repetitiveness of the GS linker might be problematic during PCR-based assembly of variable heavy and light chains, Therefore, this might bring the variation in linker length [182, 183]. There are a limited number of linkers alternative to GS linkers and different linker approaches are established for different antibody designs [138, 184-186].

In this chapter, I presented the design of the scFv fragment with two different flexible linkers, one with a non-repetitive sequence of "SPNSASHSGSAPQTSSAPGSQ", one with a repetitive sequence of $((G_3S)_4$, from full-length humanized antibody Bevacizumab, the methods used to optimize production in bacterial expression, purification, and characterization of scFv to obtain protein in higher yield and higher purities. Thereafter, I analyzed the effect of the linkers, non-repetitive and repetitive, on scFv characterization. scFv with non-repetitive linker that developed within the scope of this chapter was chosen as parental scFv for further rational design studies. Here it is concluded that scFv with non-repetitive linker showed better *in vitro* profile by having higher temperature melting point, higher monomer ratio which is desired for better long-term stability, lower immunogenicity score and lower aggregation rate. Despite scFv with non-repetitive linker showed lower antigen affinity, this property can be improved with rational design approaches. In addition to these results, biological activity of the scFvs are compared to their full-length version, Bevacizumab in zebrafish (*Danio rerio*) angiogenesis assays. Results showed that scFvs were more effective than their full-length version that might be resulted from their smaller size [173]. scFvs derived from full-length monoclonal antibody Bevacizumab, showed their potential that can be improved by further studies.

3. Investigation of scFv developability through rational design approaches : affinity-stability trade-off

3.1. Introduction

3.1.1. Antibody trade-offs

"Trade-off" can be defined as the event where two desired outcomes are compromise each other. The quality of the protein is determined by its (i) biophysical properties such as folding, stability, affinity, (ii) functional properties such as its function, specificity, solubility, and evolutionary properties such as its primary amino acid sequence and gaining new function [187]. Thus, the number of the trade-offs in proteins increases due to the higher number of desired properties at the same time. In antibody engineering field, most common trade-offs include affinity, specificity, stability, and solubility. The success of the antibodies depends on these developability properties and how the trade-offs are managed during the early development stage of the antibody. Gaining insight of the developability properties in molecular level is challenging, that might be a limit to produce optimized, effective therapeutic antibodies [188].

The generation of antibodies with increased affinity often negatively correlated with the stability [189-191]. In amino acid level, hydrophobic amino acids content of the CDRs is important for primary antigen recognition while exposed hydrophobic residues contribute to aggregation and lower stability [59, 192]. Even if affinity improvement studies generally focus on the CDRs [193, 194], several studies show that non-CDR regions can impact overall structure of the antibody, therefore affinity and the stability [195, 196]. When engineering the affinity of a molecule, it is crucial to ensure the overall stability is maintained. To achieve this, it becomes essential to perform co-screening of both stability and affinity, and to employ compensatory mutation design. These approaches are necessary to identify optimal variants that strike a balance between enhanced affinity and maintained stability [197, 198].

3.1.2. Chapter overview and publications

We rationally designed mutations on an anti-VEGF scFv that is optimized with a non-repetitive linker in chapter 2 to determine mutations to improve affinity stability and overcome the trade-off. Our focus was on the specific residues near the salt bridge, an ionic interaction formed between the residues 94 and 101 of heavy chain. This salt bridge is known to be a highly

conserved interaction that plays a crucial role in supporting the canonical structure of the HCDR3 [199, 200]. Modifying or altering this interaction might be unfavorable for both stability and affinity [201]. Here we considered the interacting residues of this salt bridge to have the potential to evaluate the affinity-stability trade-off. We chose one position from the light chain (V_L-Y49) and four positions from the heavy chain (V_H-V2, V_H-Y27, V_H-S76, V_H-V102) based on their *in silico* analysis We designed 9 mutations, recombinantly produced in *E.coli* expression system and determined their stability and affinity by differential scanning fluorimetry (DSF) and surface plasmon resonance (SPR). The mutations on V_H-V102, V_H-V102Y and V_H-V102D, showed an increase in both affinity and stability which is promising to understand the underlying reasons that alter affinity-stability properties. Then, we performed molecular dynamics simulations for gaining molecular-level understanding. We concluded that interactions near the conserved salt bridge are crucial for affinity via affecting the HCDR3 orientation towards VEGF and for stability via affecting the V_H-V_L interface. Thus, secondary interactions, that are modulated with the mutations within the scope of this chapter, play an essential part in the co-evolution of affinity-stability. Since the Vernier zone is a common region in all antibodies, implications from this chapter can be easily applied to other antibodies.

The results from this chapter were published in Biochimica et Biophysica Acta (BBA) – Proteins and Proteomics as an original research paper in 2023 [202] . Some parts of this study were selected for poster and oral presentations in congresses; "Onkolojide İz Bırakanlar Zirvesi", 14-19 Nov 2019, Antalya, Turkey – Oral presentation, V. Turkey *in vitro* Diagnostic Symposium, 19-21 Feb 2020, Izmir, Turkey – Oral presentation, European Molecular Biology Organisation (EMBO) Virtual Practical Course, Integrative modeling of biomolecular interactions, 30 May – 5 June 2021 - Poster presentation. A part of this study was also presented and published as a congress proceeding Proceeding of the 2nd International Ion Channel and Cancer Congress (IonCC2019) in the Turkish Journal of Biochemistry [176].

3.2.Material and Methods

3.2.1. Rational design analyses of mutational regions

ScFv structure without linker was extracted from Bevacizumab PDB structure, 1BJ1 [160]. Either structure or sequence of the scFv was used as input to different protein analysis webservers, Tango [203], Waltz [204], Pasta 2.0 [205], AggreScan [206], Aggrescan3D [207], CamSol sequence and CamSol structural [208, 209] that predict destabilizing, solubility-decreasing or aggregation-prone regions. The residues that were detected as problematic on at

least 5 of the 7 servers were determined as hotspot regions for further analysis. Hydrogen bonds (H-bonds), salt bridges, and hydrophobic interactions between wild-type scFv and VEGF were defined by PDBePISA-Interfaces [210] and verified by PyMOL 2.2 and molecular dynamics analysis. The interacting residues of the salt bridge between K94:D101 at the end of HCDR3 loop were analyzed by PyMOL 2.2 software. Vernier zone residues were determined according to the literature [26, 27, 211]. All the conservation data of the residues were extracted from AbYsis database [100]. Residues: V_L-Y49, V_H-V2, V_H-Y27, V_H-S76, and V_H-V102 were chosen for further mutational design. The secondary structure of the wild-type scFv and mutated scFvs were determined by using *in silico* prediction tool called SABLE [212].

3.2.2. scFv construction and expression

The anti-VEGF wild type scFv generation was detailed in previous chapter. The mutated antibodies were prepared via using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) with the primers listed in Table 3.1. The scFv mutants were transformed and expressed by using the optimized protocols within the scope of chapter 2.

Table 3. 1 Primers used for scFv mutants

	Mutant	Forward (5' to 3')	Reverse (5' to 3')
1	V_L-Y49K	ggcaaggcacccaaggttcttattaagttcacaagctcgt	acgagcttgtgaacttaataagaaccttgggtgccttgcc
2	V_L-Y49N	caacgagcttgtgaaattaataagaaccttgggtgcc	ggcacccaaggttcttattaatttcacaagctcgttg
3	V_L-Y49D	ggcacccaaggttcttattgatttcacaagctcgttg	caacgagcttgtgaaatcaataagaaccttgggtgcc
4	V_H-V2F	gccgggtagtcaggagttccagttggttgaaag	ctttcaaccaactggaactcctgactacccggc
5	V_H-Y27A	gtttaagctgtgcagcatcaggggccacatttacaaattatggtatga	tcataccataatttgtaaatgtggcccctgatgctgcacagcttaaac
6	V_H-Y27F	gctgtgcagcatcagggttcacatttacaaattatggt	accataatttgtaaatgtgaaccctgatgctgcacagc
7	V_H-S76R	ttcacttgacacttccaaaaggacagcgtacttacaaatg	catttgtaagtacgctgtccttttggaagtgtcaagtgaa
8	V_H-V102Y	tacggttcatcccattggtatttcgattattggggtcaggggac	gtcccctgaccccaataatcgaaataccaatgggatgaaccgta
9	V_H-V102D	catcccattggtatttcgatgattggggtcaggggacat	atgtcccctgaccccaatcatcgaaataccaatgggatg

3.2.3. Protein purification and characterization

Wild type and mutated scFvs were expressed and purified according to optimized protocols within the scope of chapter 2. Briefly, supernatants were collected after at high centrifugation. Protein-containing supernatants were either incubated with His-Pur Ni-NTA resin (Thermo Fisher) or loaded onto the HisTrap column and purified according to recommended commercial

protocol. Purified protein was buffer exchanged into 1X PBS (pH 7.4) through membrane filtration. After, purification proceeded with HiTrap ™ Protein L column (GE Healthcare) according to recommended commercial protocol. Protein purities were confirmed on SDS-PAGE and Western Blot. The protein amount was determined using a calculated extinction coefficient and molecular weight of scFv by Nanodrop 2000 at 280 nm.

Thermal melting points of the scFvs were determined according to optimized protocol within the scope of chapter 2. Briefly, thermal melting points were determined by DSF by using a hydrophobic dye, at ABI 7500 Fast RT-PCR. For optimal, dye and 5 µM scFv were defined and used for all studies. T_m values were defined using in-house Python scripts by utilizing Hill equation fit. Binding kinetics of the scFvs were determined by utilizing the optimized protocols in chapter 2. Briefly, VEGF was immobilized on a CM4 chip, a range of concentration of scFvs (0-100 nM) were injected onto VEGF-immobilized surface. Collected data were corrected based on VEGF included and only buffer data. Obtained sensograms from the SPR were analyzed via manufacturer's software, BiaEval 3.0. K_D, k_{on} and k_{off} were determined by fitting the sensogram curves to a 1:1 binding model.

3.2.4. Molecular dynamic and *in silico* analyses

ScFv fragments in crystal structures were prepared from the full-length antibody, Bevacizumab, crystal structure by excluding the constant regions. Mutations were introduced by using the Wizard Mutagenesis tool of PyMOL. *Homo sapiens* distributions of the residue of interest were collected from AbYsis database (www.abysis.org). For molecular dynamics (MD) simulations, MD simulations were performed under Kale Lab collaboration at IBG. Details of simulation and the analyses were given in our published article [202]. Briefly, prepared scFv structures with VEGF were solvated in water box by supplying required buffer environment. After sufficient amount of energy minimization, simulations were recorded at 310K, 1 atm, for 500 ns with 2 fs of integration time steps that resulted 5000 trajectories. For analyses, CDRs and FWs of the scFvs were determined according to Kabat numbering system [213] by using SabPred-Anarci server [214]. The contacts between variable domains were quantified as stability indicator. The contacts between scFv and VEGF were quantified as affinity indicator. Contacting residues were defined based on cutoff of 5 Å. Center of mass of the residues and/or atoms were utilized to quantify the distance between given residues and/or atoms. Root mean

square fluctuations (RMSFs) were determined utilizing the alpha carbon atoms of the proteins. The buried surface area between variable chains were quantified by extracting the solvent-accessible surface area (SASA) of the paired variable chains from the sum of the SASAs of the unpaired variable chains. Analyses and visualizations were performed by utilizing in- VMD-Python libraries Python's Numpy. Gromacs, in-house Python scripts.

3.3.Results

3.3.1. Rational behind mutational designs

The anti-VEGF scFv was previously designed and characterized with a non-repetitive linker [173]. The wild-type structure of the scFv was evaluated in different aspects to determine mutational regions for improving affinity or stability or both in favor of affinity-stability trade-off. Destabilizing regions, VEGF interaction residues, Vernier zones, and HCDR3 of the scFv were analyzed considering sequence, structure, conservation, and intramolecular interactions and the residues humanized and back-mutated during the humanization [17].

Five residues, one from variable light domain (V_L-Y49) and four from variable heavy domain (V_H-V2, V_H-Y27, V_H-S76, V_H-V102) are elected for co-evaluation of affinity and stability through secondary and tertiary interactions of the salt bridge, between K94 and D101 within the heavy chain that underlies the HCDR3 (Figure 3.1A).

V_L-Y49 is a conserved light chain framework residue. It plays role at the V_L-HCDR3 interface. This residue is included in π-π contacts with V_H-Y100E which is an HCDR3 residue. Besides, V_H-Y100E might have a hydrophobic or a anion-π contact with V_H- D101 [215]. Therefore, these interactions make an anion-π-π interaction at the interface (Figure 3.1B). For V_L-Y49, we designed three mutations, Y49N, Y49K and Y49D by considering the new formation of anion-π-amino, anion-π-cation and anion-π-anion interactions, respectively. Also, by replacing the aromatic amino acid with charged or neutral amino acids, we aimed to improve solubility that might be possible to increase stability. However, we also aimed to observe the disruptive effect of V_L-Y49D mutation due to possible of back-force between two charged amino acids, D49:D101.

V_H-V2 is one of the Vernier zone residues that is a neighbor to HCDR1 residues (V_H-S25, V_H-G26, V_H-Y27), salt bridge residue V_H-K94 and the key residue V_H-V102 within 4 Å. However, it has contact with only V_H-G26. For V_H-V2, we designed one mutation, V2F to evaluate both

its importance on overall structure and its effect as a core residue (36% solvent accessibility) in the middle of the Vernier zone residues. It was also considered to create new π-π contacts with V_H-Y27 by substitution to aromatic amino acid phenylalanine.

V_H-Y27 is another Vernier zone residue, that is a part of structural loop within the HCDR1. This residue makes important interactions within HCDR1 and plays role on loop structure. For V_H-Y27, we designed two mutations, Y27A and Y27F. Y27A is designed to evaluate the importance of aromatic side chain in local packing of the structure. The most common residue phenylalanine (48%) followed by tyrosine (18%) and phenylalanine in this position previously shown that makes extra interaction with the other Vernier residues, V_H-29 and V_H-71, forms a contiguous triad, resulted unique canonical structures of the HCDRs [27]. In addition to these, back mutation of residue 27 to phenylalanine improved antigen binding based on changing the packing between CDR and framework residues in early humanization studies [27]. We introduced the Y27F mutation to compare the two highly conserved amino acids.

V_H-S76 is located within the fourth antigen-facing non-CDR loop. This residue interacts with V_H-A24 in our structure. The most common residue asparagine (66%) followed by serine (22%). Naturally occurring asparagine in this position previously shown that interacts with V_H-Y27 and V_H-T28 (it is also threonine in our structure), stabilize the motion of the HCDR1. Naturally occurring serine in this position, interacts with V_H-S25 (it is also serine in our structure) [216]. For V_H-S76, we designed one mutation, S76R to evaluate both its importance on overall structure and its effect on stability and affinity while being very distant residue from the antigen binding region. It was also considered to create a local increased solubility by substitution to positively charged arginine.

V_H-V102 is one of the HCDR3 residues and does not interacts with VEGF. In this position, highly shared residue is tyrosine (33%) followed by valine (24%) in *Homo sapiens*. Besides, this residue is important for canonical structure of the HCDR3 by helping the stabilization of the loop. For V_H-V102, we designed two mutations, V102Y and V102D (Figure 3.1C). We introduced the V102Y mutation to compare the two highly conserved amino acids. We designed V102D mutation to evaluate possible interaction with salt bridge residue V_H-K04 and possible improvement in stability based on replacing a residue with charged residue at a solvent accessible position.

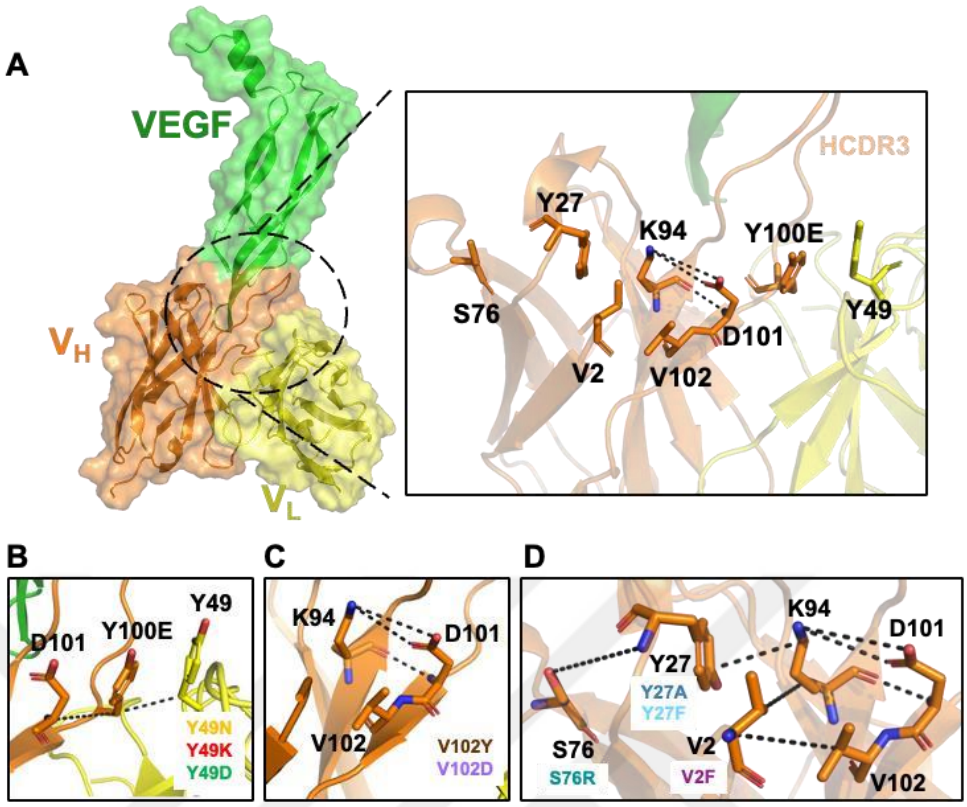


Figure 3. 1 Designed mutations near of the conserved salt bridge.

Mutations near the conserved salt bridge between K94:D101 of the V_H. A) View of the anti-VEGF scFv and antigen VEGF. B-D) Zoomed mutational landscape designed within this thesis chapter 3. B) V_L-Y49, C) V_H-V102, D) V_H-V2, V_H-Y27, V_H-S76.

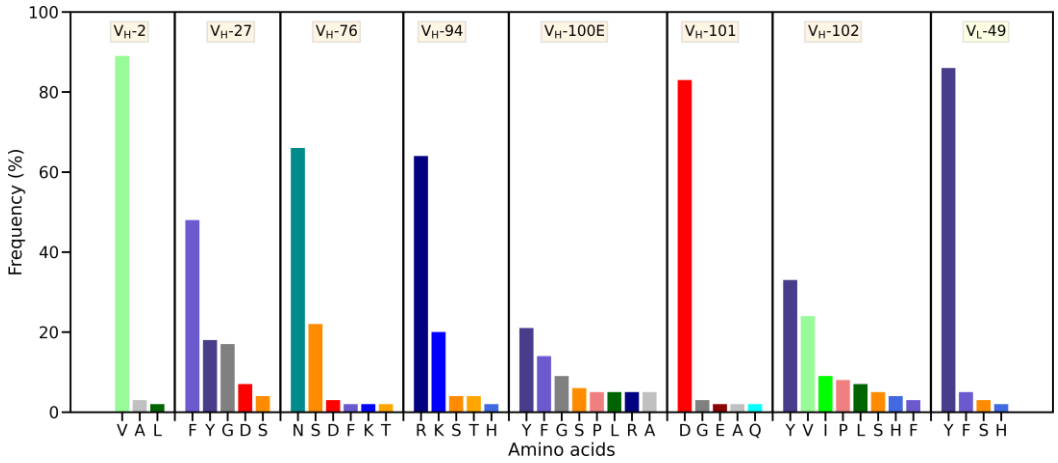


Figure 3. 2 *Homo sapiens* amino acid distributions of selected positions in this thesis

3.3.2. Affinity and stability profiles of designed scFvs

The conserved salt bridge between K94:D101, that underlies the HCDR3, play roles in stability of this loop (Figure 3.2). HCDR3 is the fundemantal binding region for most antigen interactions [217], thus we rationally designed nine mutations around this salt bridge, mostly on residues included in Vernier zone. We aimed to improve both affinity and stability with designed mutations, considering HCDR3 importance due to its variable light chain interface. As a first set of the mutational landscape, we chose residues V_H-V102 and V_L-Y49, as a second set of the mutational landscape, we chose V_H-V2, V_H-Y27 and V_H-S76 to observe the effects on affinity and stability. *In silico* prediction of secondary structures of scFvs showed that single point mutations are well tolerated (Figure 3.3).

However, when we checked the expression of mutants in supernatants, mutants V_L-Y49D, V_H-V2F could not be produced and V_H-Y27F could not be purified (Figure 3.4A). The other mutated scFv proteins were produced and purified from the supernatant in high purities (Figure 3.4B-D). Here, we designed the Y49D mutation by considering its disruptive effect that resulted in insufficient expression. On the other hand, even V_H-V2F and V_H-Y27F are the mutations to a more conserved residue for this specific position, the hydrophobic, bulky nature of the phenylalanine may have been disruptive to structure folding and stability. This resulted in the V_H-V2F mutant not being produced completely and the V_H-Y27F not being purified.

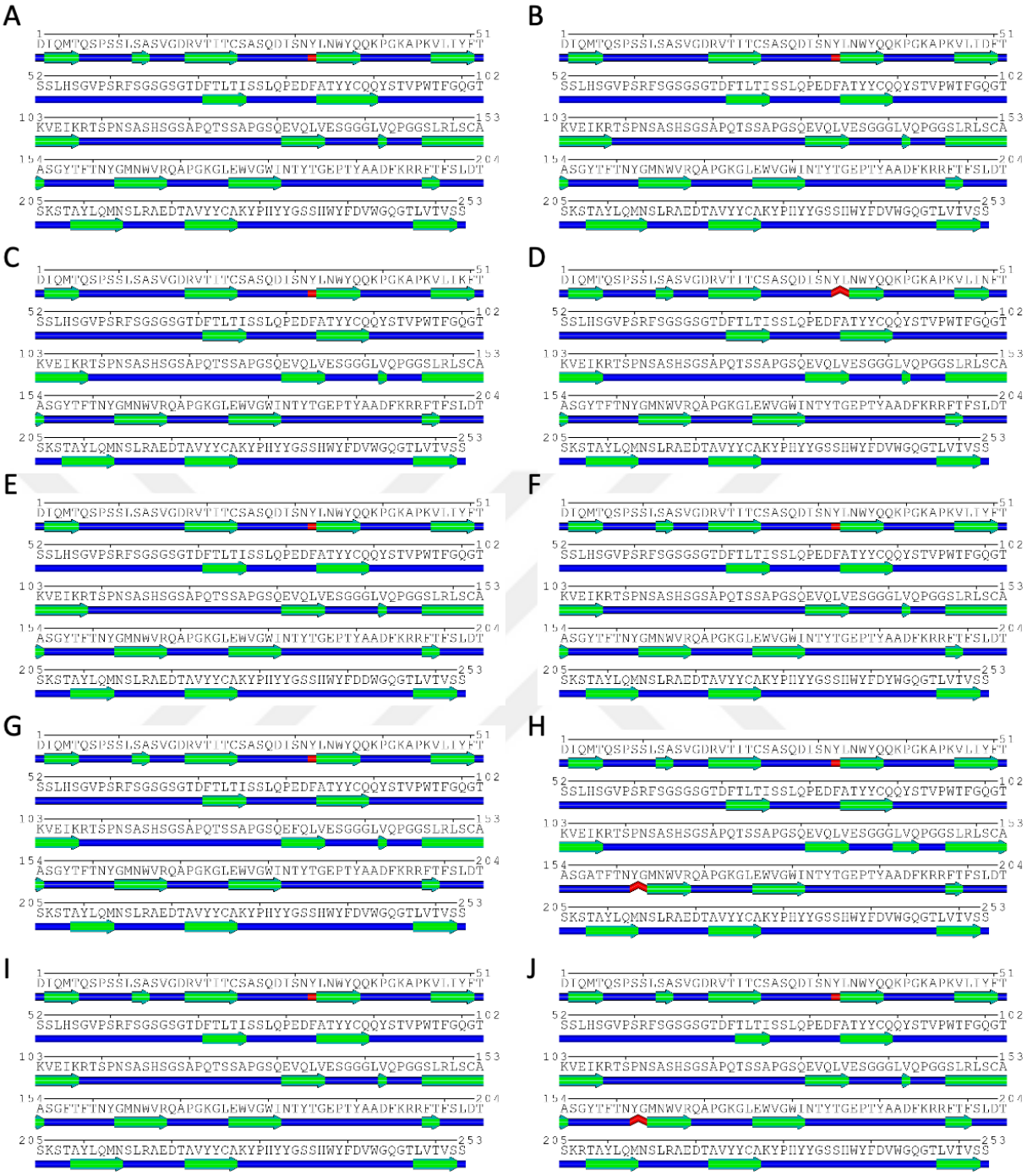


Figure 3. 3 *In silico* prediction of secondary structures wild type and mutated scFvs.

(A) WT (B) V_L-Y49D (C) V_L-Y49K (D) V_L-Y49N (E) V_H-V102D (F) V_H-V102Y (G) V_H-V2F (H) V_H-Y27A (I) V_H-Y27F (J) V_H-S76R.

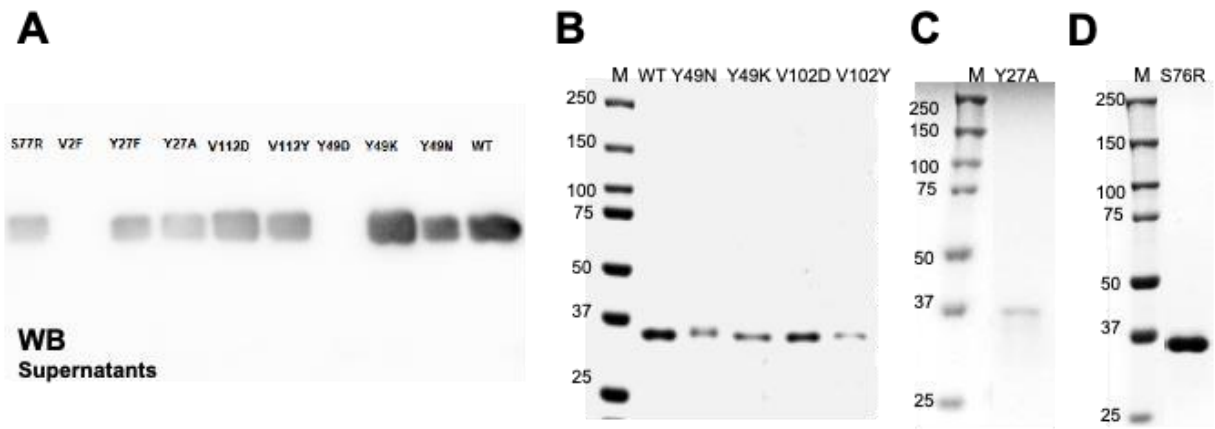


Figure 3. 4 SDS-PAGE and Western Blot (WB) analysis of scFvs

SDS-PAGE and WB analysis of scFvs after bacterial expression and purification. A) WB of expression supernatants, SDS-PAGE analysis of mutants B) WT**,** V_L-Y49N**,** V_L-Y49K**,** V_H-V102D**,** V_H-V102Y C) V_H-Y27A, D) V_H-S76R, Marker (M).

Thermal stabilities of the mutants, V_L-Y49N, V_L-Y49K, V_H-Y27A, V_H-S76R V_H-V102D, V_H-V102Y were determined 44.4 °C, 46.8 °C, 46.6 °C, 48.3 °C, 51.8 °C and 54.6 °C, respectively. It is showed that, V102Y mutation showed the highest increase in thermal melting point, V_H-V102D mutation showed the small change in thermal melting point while the other mutations showed worse thermal melting point compared to wild type scFv (Figure 3.5).

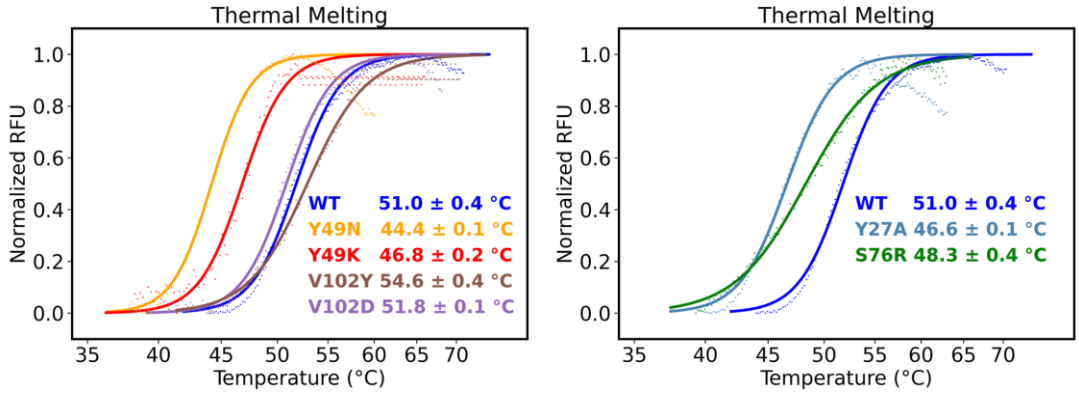
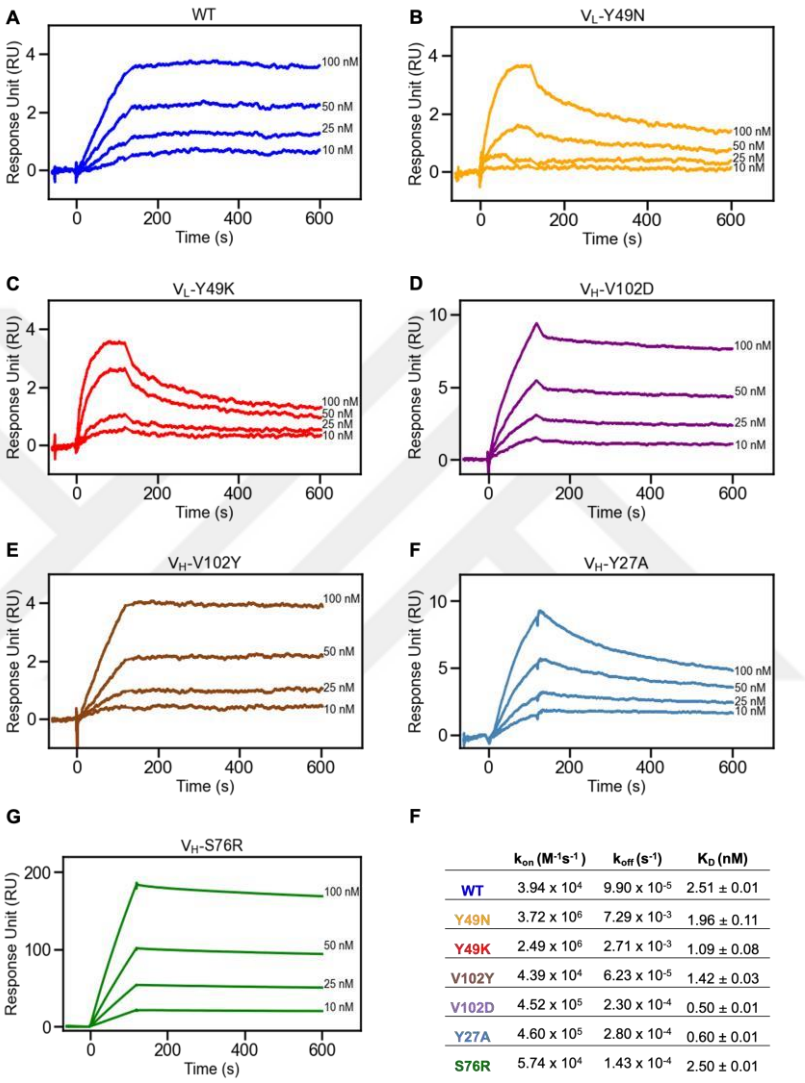


Figure 3. 5 Thermal stability profiles of the WT and mutants

Binding kinetics of the mutants were determined as 1.96 ± 0.11 nM, 1.09 ± 0.08 nM, 0.60 ± 0.01, 2.50 ± 0.01, 0.50 ± 0.01, 1.42 ± 0.03 for V_L-Y49N, V_L-Y49K, V_H-Y27A, V_H-S76R, V_H-V102D, V_H-V102Y, respectively. (Figure 3.6A-G). All the scFvs have better affinities

compared to WT, 2.51 ± 0.01 nM, except V_H-S76R. V_H-V102D showed approximately 5-fold better binding to VEGF compared to WT. Although V_L-Y49N, V_L-Y49K, V_H-Y27A, V_H-S76R mutations increased the affinity against VEGF, same improvement was not observed in their stabilities.



F

	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
WT	3.94 x 10^4	9.90 x 10^{-5}	2.51 ± 0.01
Y49N	3.72 x 10^6	7.29 x 10^{-3}	1.96 ± 0.11
Y49K	2.49 x 10^6	2.71 x 10^{-3}	1.09 ± 0.08
V102Y	4.39 x 10^4	6.23 x 10^{-5}	1.42 ± 0.03
V102D	4.52 x 10^5	2.30 x 10^{-4}	0.50 ± 0.01
Y27A	4.60 x 10^5	2.80 x 10^{-4}	0.60 ± 0.01
S76R	5.74 x 10^4	1.43 x 10^{-4}	2.50 ± 0.01

Figure 3. 6 Experimental affinity profiles of wild type and mutated scFvs.

SPR profiles of all mutants **(**A) WT (B) V_L-Y49N (C) V_L-Y49K (D) V_H-V102D (E) V_H-V102Y (F) V_H-Y27A (G) V_H-S76R. F) K_D (nM), k_{on} ($M^{-1}s^{-1}$), k_{off} (s^{-1}) values of the WT and mutants.

3.3.3. Molecular dynamic analyses

Based on their experimental results, V_H-V102 mutants are promising to overcome antibody affinity-stability trade-offs. Also, V_L-Y49 mutants are in the core of the HCDR3- V_L interface.

We further chose V_H-V102 and V_L-Y49 mutations to evaluate the effects of secondary interactions of these residues on their local interactions.

Here, we first calculated the RMSF values of the simulated scFv structures. The RMSF values stayed in a range of 6 to 10 Å. Therefore, we concluded scFv and VEGF maintained their binding complex, overall flexibility of the antibody-antigen complex remained unchanged after introducing the selected mutations (Figure 3.7).

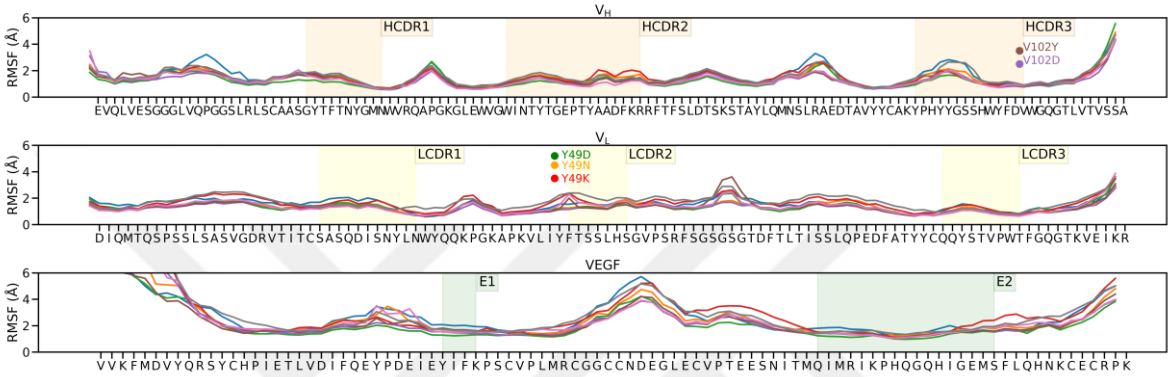


Figure 3. 7 Flexibility profile of the scFvs during the MD simulations

After that, we quantified the mean total contacts between heavy and light chain considering stability indicator and the mean total contacts between scFv and VEGF considering affinity indicator. Here, experimental results were supported by calculated data coming from MD trajectories. V102Y and V102D mutations gained both stability and affinity contacts. Y49K and Y49N mutations only gained affinity contacts while losing the stability contacts in the interface. We observed the highest stability and affinity contacts decrease in Y49D mutations explains that this mutation could not be produced and characterized recombinantly. At the end, each mutation fell into the same window of the chart according to both experimental data and simulated data. Therefore, we concluded data coming from MD simulations can be used for further analysis to gain insight for affinity-stability trade-off due to supporting the experimental data (Figure 3.8).

Since, mutations on Y49 showed the stability decrease, we utilized the mutations on V102 as rescue mutations. Here we combined the Y49N with V102Y and V102D in the MD simulations. Even both double mutations designs gained antigen contacts and might show better affinity profile, they lost the stability contacts with light chain. Even they showed lower stability

contacts than the Y49D mutation, the one could not be produced recombinantly. Therefore, those double mutations were not taken for further experimental production and characterization process.

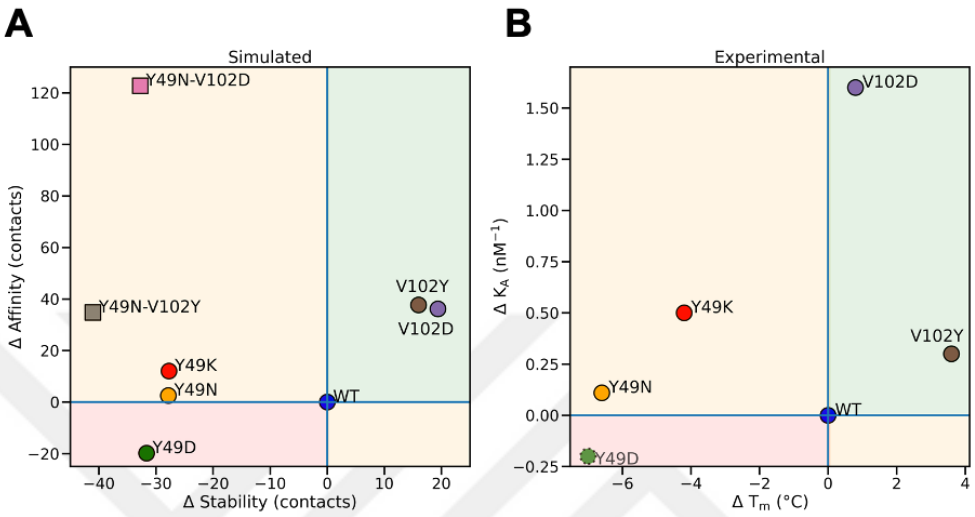


Figure 3. 8 Affinity-stability differences of scFv mutations compared to scFv wild type.

The green window indicates increase in both affinity and stability, orange window indicates the decrease in affinity and increase in stability or, increase in affinity and decrese in stability, red window indicates decrease in both affinity and stability. A) Contact count difference of mutations coming from MD trajectories. B) Experimental value difference of mutations coming from wet-lab analyses.

We further aimed to explain these affinity and stability changes in molecular level. In pairwise contact analysis, we observed changes in HCDR3 contacts for specific regions that might lead to affinity or stability increase/decrease (**Figure 3.9**). Since all the mutations led to gain in affinity contacts with VEGF, we observed increase in contacts between HCDR3 and antigen at the first 5-6 residues of the HCDR3. On the other hand, stability decreased mutations showed decrease in contacts between HCDR3 and light chain interface at the last 4-5 residues of the HCDR3 (**Figure 3.9**). Most importantly, highest gain of contacts was observed in mutations with highest affinity increase V102Y and V102D. We concluded that HCDR3 loop of our scFv utilize the loop at its two faces (between 95-100B), VEGF interacting face, taking role in antigen binding and effective in affinity, light chain interacting face (between 100B-102), taking role in

interface interactions and effective in stability. Hence, we emphasized that the HCDR3 plays a crucial role not only in affinity but also in stability.

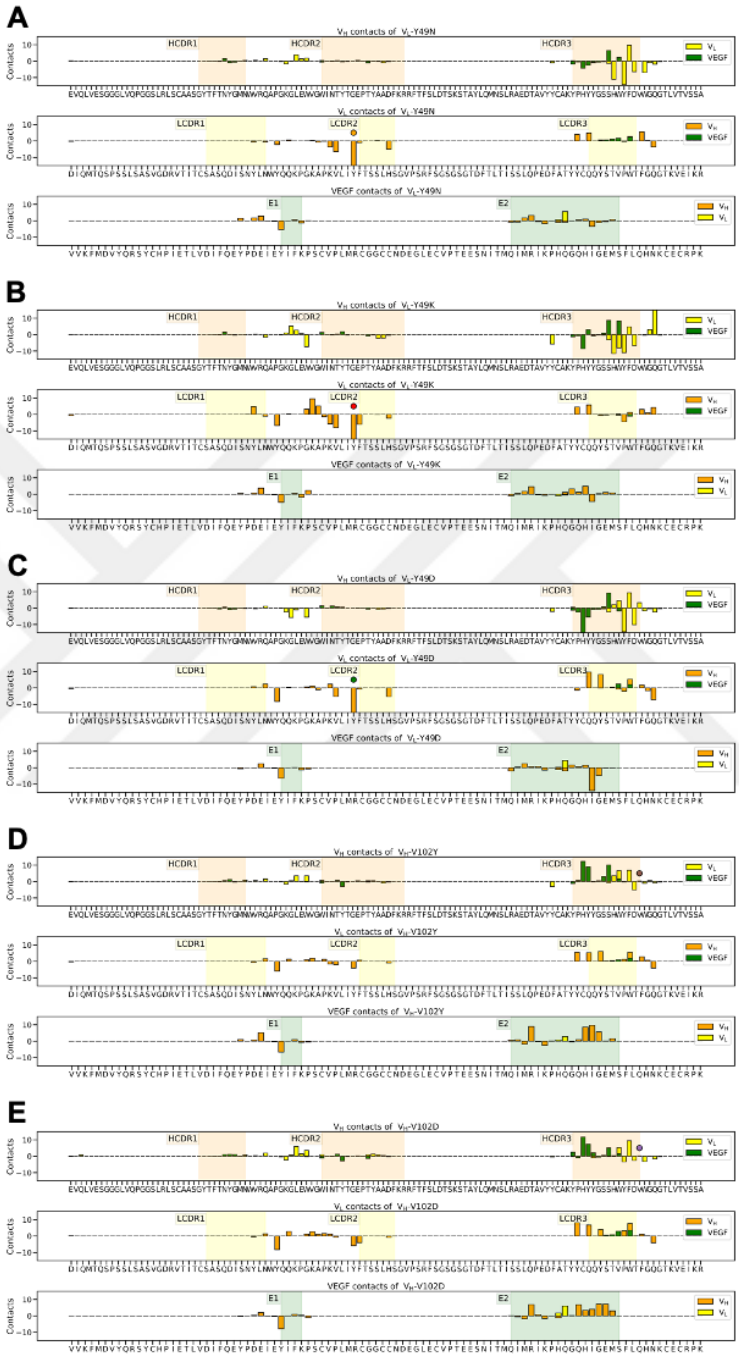


Figure 3. 9 Contact count differences of scFv mutations compared to scFv wild type.
Heavy chain contacts, light chain contacts, and antigen contacts are colored orange, yellow and green, respectively. A) V_L-Y49N, B) V_L-Y49K, C) V_L-Y49D, D) V_H-V102Y, E) V_H-V102D. The top panel represents the light chain and antigen contacts of the heavy chain. The middle

panel represents the heavy chain and antigen contacts of the light chain. The bottom panel represents the heavy and light chain contacts of antigen.

Then, we quantified the buried surface area between the variable chains to express the change in stability. Since packing between light and heavy chains effects the overall stability [218, 219], the buried surface are between those chains should be stayed similar to wild type. As expected, stability-decreased mutations showed worse buried surface areas than the wild type's which means heavy and light chain packing changed due to introduced mutations and lost contacts. These contacts lose made the variable chains distant from each other, some of the interfacing areas became solvent accessible which might lead to stability decrease (**Figure 3.10)**. On the other hand, buried surface area value of stability increased mutations remained similar or higher than the buried surface area of wild type. Here we concluded that stability-decreased mutations affected the interfacing area unfavourably between variable chains.

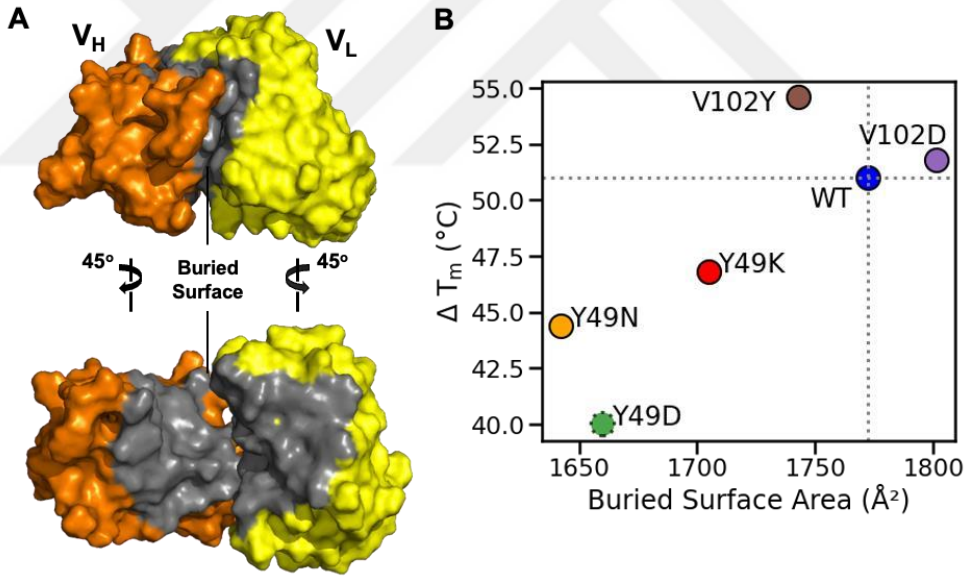


Figure 3. 10 Buried surface area of scFvs between variable chains.

A) Crystal structure representation of the quantified BSA (gray area). B) Calculated mean of BSA of each mutation plotted with T_m points.

Further, we aimed to explore the molecular level interactions of the mutations and compared to wild type. Here, when the important anion-π-π interactions between D101-Y100E-Y49 modulated with Y49 mutations, the distance between these residues were increased or disrupted with Y49 mutations (**Figure 3.11A**). However, the distance between these residues remained

similar to scFv wild type version with the V102 mutations (**Figure 3.11C).** We also checked the other HCDR2 – light chain interfacing residues and there was drastic change in other residues at the interface. We concluded that the π-π interaction between Y49-100E was crucial for heavy – light chain interface, this stacking is critical for contribution of HCDR3 to overall scFv stability. Most importantly, we observed an extra salt bridge that formed after the introduction of V102D mutations. Since it has been showed that one charged amino acid can have more than one ionic interaction with other amino acids [17, 220], the salt bridge residue K94 not only maintained its parental ionic interaction but also gained an extra ionic interaction with the replaced aspartate with valine on position 102 (**Figure 3.11B, Figure 3.11D**). We concluded that triple ionic interactions between the positions K94-D101-D102 might stabilize the HCDR3 better and affect the affinity favorably. At the end, this mutation was resulted with five times better binding affinity compared to wild-type.

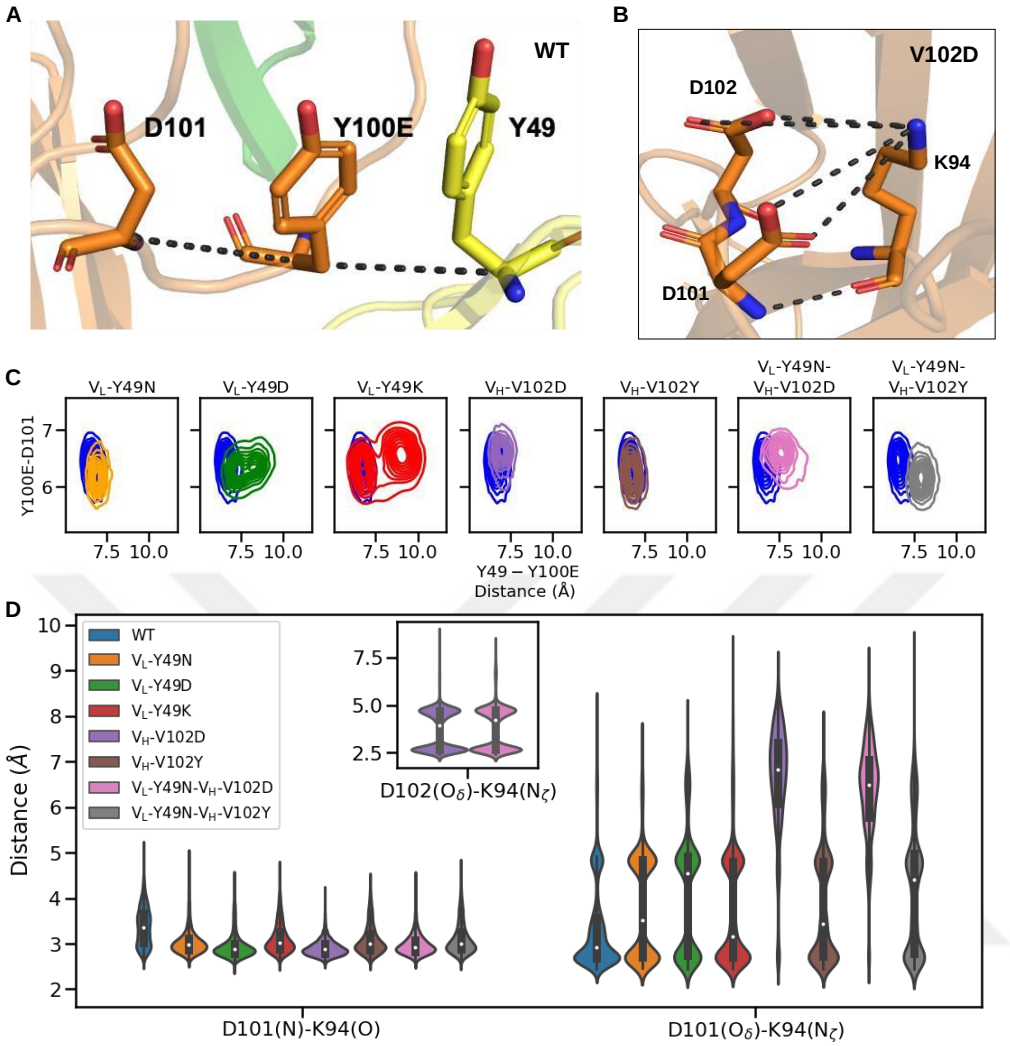


Figure 3. 11 scFv affinity and stability interactions at molecular level understanding.

A) The schematic representation of anion-π-π interaction between D101-Y100E-Y49 within HCDR3 – light chain interface. **B)** The schematic representation of the parental salt bridge between K94-D101 and the triple salt bridge between K94-D101-D102. **C)** Distances between alpha carbon atoms of residues Y49-Y100E and residues Y100E-V_H-D101 residues. **D)** Distances between the interacting parental salt bridge atoms of K94 and D101. Inset shows the distance between interacting newly formed salt bridge atoms of K94 and D102.

Next, we observed an orientation change in HCDR3 abiding by antigen in mutated scFvs compared to wild-type scFv (**Figure 3.13A**). We quantified the altered orientation due to mutations by calculating the angle between HCDR3 and VEGF where the center of mass of the variable domains, antigen, and middle of HCDR3 were utilized **(Figure 3.13B).** Here,

quantified angles showed that all the scFvs with mutations had smaller angles which means they gained closer orientation to the antigen. This angle changes to smaller ones might explain the affinity increase in all mutations compared to the wild-type. Most importantly, the smallest angle was observed within the scFv with mutation V102D which was supported by the experimental data by having the 5 times more binding affinity compared to wild-type. We concluded that introduced mutations contributed to affinity increase by altering the orientation and/or stabilizing the HCDR3 loop structure (**Figure 3.13B)**.

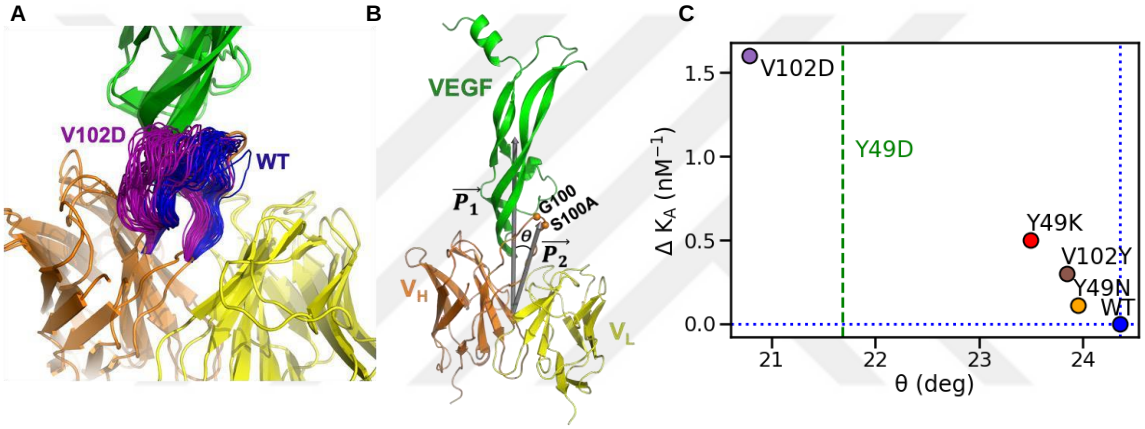


Figure 3. 12 The altered angle between HCDR3 and VEGF

A) Crystal structure representation of the HCDR3 loop orientations of wild type and mutation V102D. **B)** The angle between HCDR3 and VEGF where the center of mass of the variable domains, antigen, and middle of HCDR3 were utilized. **C)** Quantified angle changes of mutation introduced scFvs compared to wild-type scFv Angle data calculated from the MD trajectories.

3.4. Discussion

Antibodies and its different formats are generated for various applications and developability engineering efforts might bring trade-offs because of the complex intramolecular interactions of the antibodies [187]. Improving affinity during the antibody development process is desired but these efforts might affect the other properties Affinity-stability is the most shared trade-off of the antibody engineering efforts [188, 197, 221]. HCDR3 could be the key region for both

affinity and stability due to both being in antigen-binding interactions and being at the interface of variable domains.

In this chapter, even though HCDR3 is the first choice to modulate affinity in engineering approaches, no mutations were introduced directly into the HCDR region. Instead, we considered the residues which are either Vernier zone or close to Vernier zone residues to alter the affinity and stability profile of the generated scFv. For that, we designed nine mutations and obtained affinity and stability increases in two mutations on heavy chain residue V102, V102Y, and V102D. This residue not only permitted mutations that changed the orientation of the HCDR3 to a more favorable position toward antigen but also maintained or gained interactions at the light chain interface resulted in an increase in stability. Results from this chapter suggested that HCDR3 loop is not only an affecting factor for affinity but also an affecting factor for stability. In addition to these, the very conserved ionic interaction between heavy chain K94-D101 residues underlies the HCDR3 loop. Whenever, an extra salt bridge is created next to the parental one, both affinity and stability increased. A more-depth analyses showed that complex salt bridge resulted from V102D mutation made HCDR3 lean on VEGF further which could explain its affinity increase.

Here, we demonstrate the importance of Vernier zone residues for antibody engineering efforts although they are underestimated in current literature.

4. A proof-of-concept study on antibody specificity modulation : mono- to dual- specificity

4.1.Introduction

4.1.1. Yeast display systems

The antibody engineering techniques has significantly accelerated the progress of therapeutic antibody development. Various protein engineering approaches can be employed to enhance crucial properties of antibody fragments, including affinity, specificity, and stability. [222]. Affinity improvement and/or maturation is an essential step for therapeutic antibody development because it determines biological activity and clinical efficacy [222]. One of the affinity maturation approaches for higher affinities to target antigens is the modulation of antigen-binding regions of antibodies by display techniques [96]. Display techniques (e.g., phage display, mammalian, and yeast surface display) are the most powerful screening

techniques for that kind of purpose, as they enable very rapid screening of antibodies with the desired characteristics from synthetic/semi-synthetic or natural antibody libraries [94].

Among the display systems, yeast surface display is a well-founded directed evolution strategy for the discovery and development of antibody fragments with advantages. Within this method, the proteins being showcased undergo folding in the endoplasmic reticulum of eukaryotic yeast cells, where they can take advantage of quality-control mechanisms of the yeasts [87]. It requires less time/cost compared to other eukaryotic systems [223]. Also, it enables quantitative screening via Fluorescence-Activated Cell Sorting (FACS), allowing for direct observation of the equilibrium activity and statistics of the sample throughout the screening process [224, 225]. Methylotrophic yeast *P.pastoris* is the most preferred yeast species for recombinant protein production due to higher cell density, higher protein yield and less glycosylation [226]. There are many antibody display studies on *Pichia pastoris* with different cell wall proteins adapted from other species such as agglutinin proteins (*S. cerevisiae* cell wall agglutinin protein 1, Sag1) [227, 228]. There is also a cell wall anchor protein "protein with internal repeats of *P.pastoris*" (PpPIR1) [229]. PpPIR1 system of *P.pastoris* has only been used for non-antibody proof-of-concept display studies [230, 231]. This is the first study testing the PpPIR1 system for scFv antibody fragment according to our literature research.

4.1.2. Bispecific antibodies

The many of different and complex biological pathways are associated with tumor growth that often pose challenges for the success of the treatment while using mono-specific targeting agents. Lately, the development of different targeting formats of antibodies, such as bispecific antibodies, has emerged as a strategy to enhance therapeutic efficacy [232]. Bispecific antibodies can bind to two distinct antigens [233]. Presently, there are 14 bispecific antibodies out of the 180 approved antibodies, and this number is increasing based on their clinical success [234]. The most preferred approach to generate bispecifics involves two different antibodies targeting separate antigens, each located on separate arms of the antibody structure [235] **(Figure 4.1A)**. But then, dual specifics, might be referred to as "Two-in-One", possess the capability to bind to two antigens individually using the same antigen-binding site [236] **(Figure 4.1B)**.

Antibody specificity modulation have an immense potential for discovering more efficient antibodies for therapeutic approaches. Generally, HCDRs are the main driver in antigen

binding. Hence, LCDRs might be possible paratope for second antigen binding. Very first proof-of-concept engineering approaches implemented on mono-specific antibodies to obtain dual specifics [61, 236-241]. Bostrom *et al.* utilized a LCDR library of anti-Human Epidermal Growth Factor Receptor 2 (HER2) antibody, Herceptin, discovered dual specific antibodies with different affinities against VEGF and HER2 [237]. Their study demonstrated that introducing specific mutations in the LCDRs was adequate to achieve dual-specific antibodies capable of binding to both HER2 and VEGF. This pioneering technique provided evidence that monospecifics can be engineered to acquire dual- and/or multi-specificity by introducing mutations in the avaiable CDRs.

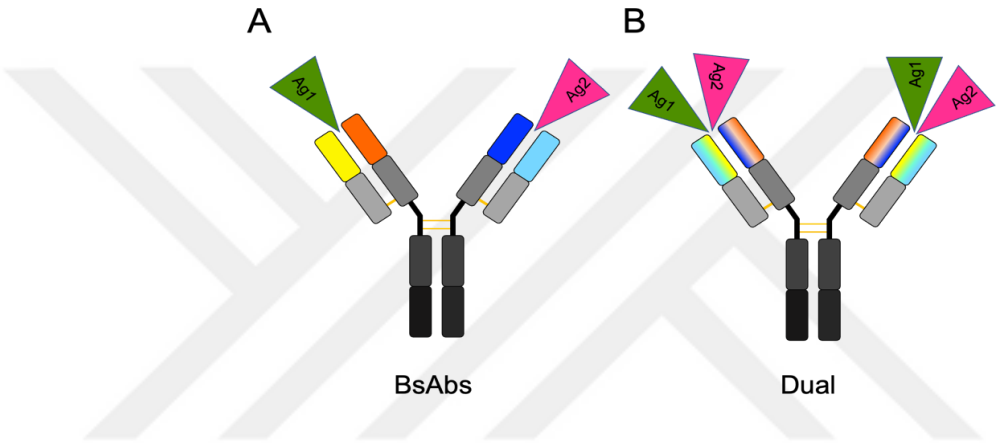


Figure 4. 1 Antibody targeting formats.

A) Bispecific antibody, two different variable regions target two different antigens. B) Dual-specific antibody, same variable region targets two different antigens. Antigens were shortened as Ag1 (green) and Ag2 (pink). Constant regions were indicated in different shades of gray. Variable heavy chains were indicated in orange and dark blue, variable light chains were indicated in yellow and light blue.

Efforts to modulate antibody specificity and/or affinity predominantly focus on the complementarity-determining region (CDR) regions. The primary distinctions among variable domains of antibodies lie in the properties of CDR loops such as amino acid content, and. In contrast, the non-CDR regions are typically conserved and exhibit a high degree of structural similarity, formed by several core β-sheet structures [242]. While some studies highlight the significance of non-CDR regions in biophysical properties [188] and humanization [28, 243], their contributions to affinity/specificity are often underestimated [188, 244]. The Vernier zone,

a critical non-CDR region, has the potential to influence the canonical structure of CDR loops [26, 27]. However, no study has yet investigated the relationship between Vernier zone and antibody specificity [242]. While Vernier zone is commonly engineered/back-mutated during humanization applications to restore affinity, the features of Vernier zone have potential to impact various antibody properties, such as specificity.

4.1.3. Dual blockade of VEGF and PD-L1

Antibodies used in cancer treatment can be used alone or in combination for targeted therapy approaches. Programmed Death-Ligand-1 (PD-L1) has emerged as a crucial protein in the cancer therapies and diagnostics. PD-L1, expressed by tumor cells, is presented on the surface, binds to the programmed-death-protein-1 (PD-1) on T cells in the immune system, causing immunosuppression against cancer [245]. Anti PD-L1 therapy is emerging as an effective and successful treatment for cancers with high PD-L1 expression. It has been reported that the combination of the anti-angiogenesis and anti-immunosuppression approaches in cancer treatment increases drug efficacy in various cancers [246]. The crosstalk between immunosuppression and angiogenesis indicates that the remodeling of the cancer vessel network can enhance the effectiveness of cancer immunotherapies. Dual targeting of PD-L1 and VEGF has great potential for a combinatorial treatment approach [247, 248].

Currently there are 9 approved antibodies (out of 180) and 4 antibodies (out of 138) in late-stage clinical studies that blocks the PD-L1/PD-1 interaction for cancer therapies (Table 5.1, extracted from Antibody Society [158]).

Table 4. 1 PD-L1 blocking antibodies approved and in late-stage clinical studies

Approved antibodies				
Antibody	Target	Format	Specificity	Therapeutic Area
Tagitanlimab	PD-L1	Full-length	Monospecific	Cancer
Sugemalimab		Full-length	Monospecific	Cancer
Socazolimab		Full-length	Monospecific	Cancer
Envafolimab		sdAb-Fc	Monospecific	Cancer
Durvalumab		Full-length	Monospecific	Cancer
Cosibelimab		Full-length	Monospecific	Cancer
Avelumab		Full-length	Monospecific	Cancer

Atezolizumab		Full-length	Monospecific	Cancer
Adebrelimab		Full-length	Monospecific	Cancer
Antibodies in late-stage clinical studies				
Antibody	Target	Format	Specificity	Therapeutic Area
Retlirafusp alfa	PD-L1	Full-length (PD-1 fusion)	Bispecific (TGF-β)	Cancer
Erfonrilimab		sdAb-Fc *	Bispecific (CTLA4)	Cancer
Bintrafusp alfa		Full-length	Bispecific (TGF-β)	Cancer
TQB2450		Full-length	Monospecific	Cancer

*Erfonrilimab is a novel fusion antibody design, two anti-PD-L1 single domain antibodies are positioned in place of variable domains, two anti-CTLA4 single domain antibodies are positioned in place of C_H1-C_L. This sdAb design is fused to Fc region.

4.1.4. Chapter overview

High specificity is one of the important determinants of an antibody's success. Mono-specificity is usually desired to prevent off-target binding but controlled multi-specificity could be advantageous in treatments requiring more than one antigen target. There is a promising multi-specific format called dual-specific antibodies which can exhibit binding to two distinct antigens while utilizing the same complementarity determining regions. Dual-specific binding is generally modulated from mono-specific binders via diversification of the CDRs, but roles of non-CDR regions in antibody specificity is underrepresented in the literature.

One of important non-CDR regions is Vernier zone. Although its role in affinity is documented, its effect on specificity is not clear. We previously showed that interaction between a non-CDR antigen facing loop including Vernier zone residues (Light-Vernier-4, LV4) and CDRs might be key to gain dual-specificity, detailed in Chapter 4. Here, we hypothesized that diversification of LV4 loop of a mono-specific antibody can lead to a dual-specific binding without any direct mutations in CDRs.

In this chapter, mono-specific single-chain antibody variable fragment against Vascular Endothelial Growth Factor (VEGF) was used as a template. Two residues of the LV4 loop were diversified, and the corresponding library was displayed on the surface of *Pichia pastoris* with *S. cerevisiae* cell wall protein Sag1. Screening was performed for a second antigen, Programmed Death-Ligand-1(PD-L1), to obtain a dual-specific antibody. In this context, Sag1 display system was optimized for scFv antibody fragment, prepared scFv library was sorted to obtain PD-L1 binding scFvs. One enriched clone, with NQ motif in the corresponding residues,

showed PD-L1 binding in the therapeutic range, ~50 nM and preserved its VEGF binding after mutations. Our study is a novel approach to modulate specificity of the antibody that is not required large libraries including CDRs.

4.2. Material and Methods

4.2.1. Generation of wild type scFv and scFv library display plasmids

The anti-VEGF scFv gene was amplified from pET-17b bacterial expression plasmid by adding either SfiI/PacI or EcoRI/ApaI restriction sites. For the PpPIR1 based display plasmid construction, the scFv gene with N-terminal EcoRI and C-terminal ApaI restriction sites was ligated with the orientation of PpPIR1-V_L-linker-V_H-Myc-6xHis into yeast surface display plasmid. For the Sag1 based display plasmid construction, the scFv gene with N-terminal SfiI and C-terminal PacI restriction sites was ligated with the orientation of Flag-V_L-linker-V_H-V5-Sag1 into yeast surface display plasmid (*P. pastoris* pPSDZeoSfiIPacI-FLAGV5-AOX1 surface display vector). Then, either DH5a or TOP10 electrocompetent *E.coli* cells were transformed with these ligation products and isolated plasmid DNAs were verified by DNA sequencing (Eurofins Genomics). KOD DNA Polymerase (TakaraBio) or Phusion DNA Polymerase (NEB) were used for amplification. All restriction enzymes and T4 DNA ligase were provided from NEB. Low salted LB-Broth medium (10g/L tryptone, 5 g/L yeast extract, 5g/L sodium chloride) and low salted LB-Agar plate (LB-Broth content, 15 g/L agar) were used for bacterial growth. 20 μg/mL zeocine was used for antibiotic selection of positive clones. Promega PureYield Plasmid System or Macherey-Nagel NucleoSpin Plasmid System were used for plasmid isolation.

The two residues of fourth loop on light chain (residues 68 and 69) were determined for specificity modulation. A library was designed by replacing each position with 20 different amino acids, the total library size was determined as $4x10^2$. The scFv gene library was prepared by overlapping PCR with degenerative primers (Overlap extension PCR). Overlapping fragment 1 was 242 bp amplified by adding SfiI restriction site at N-terminus and degenerative codons at selected positions with the scFv-FW-SfiI primer and Degenerative-2aa-RV primer. Overlapping fragment 2 was 561 bp amplified by adding PacI restriction site at C-terminus with the scFv-VL71-FW primer and scFv-RV-PacI primer. For the assemble of overlapping fragments, first 3 cycles of the PCR was amplified without primers, relied on the overlapping sequences generated in the first part. Both fragments were presented in the PCR reaction in

equal amount (75-125 ng). After the 3 cycles without primers, scFv-FW-SfiI and scFv-RV-PacI primers were added into reaction, assemble fragment was amplified for 25 cycles. Assemble fragment (788 bp) was identified on a 1% agarose TAE gel containing SYBRSafe DNA Gel Stain (ThermoScientific S33102).

The fragment library prepared via overlap PCR were cloned in Sag1 surface display vector (P. pastoris pPSDZeoSfiIPacI-FLAGV5-AOX1) described in the section 5.2.1. Enough vector and library fragment were digested with SfiI and PacI, vector was purified from agarose gel and insert was purified with the NucleoSpin cleanup kit. Prepared vector and insert were ligated in a 1:3 molar ratio with T4 DNA ligase (NEB) using temperature cycle ligation incubation (TCL) [249]. *E.coli* TOP10 cells were freshly prepared for electroporation. Electroporation was performed in several of pre-chilled 2 mm electroporation cuvettes. 40-50 µl electrocompetent cells were incubated with 2 µl of ligation reaction per cuvette. Electroporation was performed at 2.5 kV for 4-5mscec. Cells were pooled and recovered in SOC medium (5 g/l yeast extract, 20 g/l tryptone, 0.5 g/l NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM glucose) for 1 h at 37 °C. Then recovered cultures are plated on low salted LB agar plates containing 20 ug/mL zeocine. To assess library diversity, a serial dilution of the recovered cells was plated. Following overnight incubation at 37 °C, all colonies were collected from the agar plates and combined. The plasmid library was subsequently isolated from cells using NucleoBond Xtra Midi preps (Macherey-Nagel) and eluted in TrisHCl pH 8.5. Plasmid DNAs were isolated from randomly picked six colonies in the library and diversity were verified by DNA sequencing (Eurofins Genomics).

4.2.2. *P.pastoris* transformation and library generation

The yeast surface display of antibody fragments was performed using the Pichia pastoris GS115 strain. Optimized P. pastoris transformation procedure was followed [250]. Briefly, a fresh colony of GS115 was inoculated in 5 mL YPD (1% yeast extract, 2% peptone, 2% glucose monohydrate) and incubated overnight at 28°C, 250 rpm. Overnight pre-culture has inoculated into 250 mL of YPD. Culture was grown to 1.5 OD_{600} nm at 28°C and 250 rpm. Then cells were harvested by centrifugation at 1500-2000 g for 5 min at 4°C. Cells were resuspended in LiAc/DTT solution and incubated for 30-45 mins at room temperature at 100 rpm. Then cells were harvested washed twice with ice cold 1 M sorbitol. Collected cells were resuspended in 1

M ice cold sorbitol and stored the cells on ice until electroporation or flash-freezed for further use. Plasmids were digested with *Pme*I for linearization of the plasmid library in the *AOX1* promotor. Linearized plasmids were desalted by a PCR purification kit (Promega). 100-150 ng of linearized plasmid DNA was mixed with 80-100 µL of competent yeast cells, incubated on ice for 5 minutes. Electroporation was performed in a pre-chilled 2 mm electroporation cuvette at 1.5kV for 3msec. Ice cold 1 M sorbitol or YPD was immediately added after pulsing the cells. After 3-6 hours recovery, cells were spreaded on YPD agar plates containing 20 ug/mL zeocine. Plates were incubated at 28°C and 250 rpm for 2-3 days. Single colonies were chosen for overnight growth in YPD medium, then 15% glycerol stocks were prepared for further studies.

The yeast surface display of antibody fragments was carried out utilizing the Pichia pastoris GS115 strain. Optimized P. pastoris transformation procedure was followed as described above. Freshly prepared electrocompetent GS115 cells were used for transformation. A mixture of 100-150 ng of linearized plasmid DNA and 80-100 µL of competent cells was prepared and incubated on ice for 5 minutes. Electroporation was performed in multiple pre-chilled 2 mm electroporation cuvette at 1.5kV for 3msec. YPD was immediately added after pulsing the cells. Following a 6-hour recovery period, a serial dilution of the retrieved cells was plated to assess library diversity. The remaining transformed cells were then subjected to liquid selection by inoculating them into YPB broth at a ratio of 1/25, supplemented with 20 µg/mL zeocin. The culture was subsequently incubated at 28°C and 250 rpm for 24 hours. Then 15% glycerol stocks were prepared.

4.2.3. Flow cytometry analyses

GS115 cells transformed with the expression constructs were inoculated into 5 mL of YPD medium and incubated overnight (~16 - 24 h) at 30°C, 200 rpm. The overnight culture was then used to inoculate 15 mL of BMGY medium with a starting optical density at 600 nm (OD600) of 0.1. The culture was grown at 30°C until the OD600 reached a range of 6-10. Then , the cells were harvested by centrifugation at 3000 × g for 10 minutes at room temperature and resuspended in 15 mL of BMMY medium. The culture was supplemented with 1% methanol (final concentration) at 12-hour itime points during the induction period. After 24 or 48 hours of growth, the cells were harvested by centrifugation at 1500 × g for 5 minutes at 4°C.

Harvested cells were washed twice with washing buffer (1X PBS Buffer, 2 mM EDTA, protease inhibitor, pH 7.4), resuspended in cold washing buffer. 100 uL of OD_{600} 1-2/mL cells were pelleted and resuspended in 100 uL staining buffer (1X PBS Buffer, 1% BSA, 2 mM EDTA, protease inhibitor, pH 7.4). Cells were incubated either with 10 nM biotinylated VEGF [SinoBiological 11066-H27H-B] for 2 hours or with primary antibody of interest for 1 hour. 1/100 mouse anti myc-tag antibody (ProteinTech 67447-1-Ig) or 1/100 mouse anti-his tag antibody (ProteinTech 66005-1-Ig) were used for PpPIR1 display system. 1/100 mouse anti-flag tag antibody or 1/200 anti-flag rabbit antibody (SigmaAldrich F7425) was used for Sag1 display system. Cells were pelleted and washed twice with staining buffer. Cells were pelleted in 100 uL staning buffer and incubated with secondary antibody/reagent of interest. 1/200 anti-mouse Alexa Fluor 488 conjugated antibody (ProteinTech SA00013-1) was used for primary mouse antibody incubated cells. 1/500 anti-rabbit Alexa Fluor 488 conjugated antibody (LifeTech A11008) was used for primary rabbit antibody incubated cells. 1/400 Streptavidin-PE (Pharmingen 554061) or 1/200 Streptavidin-Alexa Fluor 568 (ThermoScientific, S11226) was used to detect VEGF binding. Following the incubation with secondary reagents, the cells were subjected to washing with 200 µL of ice-cold staining buffer. Subsequently, the cells were resuspended in 150 µL of ice-cold staining buffer for flow cytometric analysis. 10000 events at minimum were recorded per sample using a BD FACS Melody flow cytometry cell analyzer.

4.2.4. Fluorescence activated cell sorting (FACS) analyses

The protocol described in section 5.2.4 was followed for sorting preparation. After 24h methanol induction of AOX1 promoter the cells were harvested at 1500 g for 5 minutes at 4°C. Harvested cells were washed twice with washing buffer (1X PBS Buffer, 2 mM EDTA, protease inhibitor, pH 7.4), resuspended in cold washing buffer. 100 uL of OD_{600} 2/mL cells were pelleted and resuspended in 100 uL staining buffer (1X PBS Buffer, 1% BSA, 2 mM EDTA, protease inhibitor, pH 7.4). Cells were incubated with biotinylated PD-L1 (Sinobiological, 10084-H08H-B) in a range of concentrations (100 nM to 1 nM, decreasing in each sorting) for 2 hours and 1/200 anti-flag mouse antibody was added after 1 hour incubation of the antigen. Cells were pelleted and washed twice with staining buffer. Cells were pelleted in 100 uL staning buffer and incubated with 1/200 anti-mouse Alexa Fluor 488 conjugated antibody (ProteinTech SA00013-1), 1/200 Streptavidin-Alexa Fluor 568 (ThermoScientific, S11226) and 1/500 DAPI (ThermoScientific D1306). After the incubation with secondary reagents, cells were washed

with 200 uL ice-cold staining buffer and resuspended in 150 uL ice-cold staining buffer for FACS. 10000 events at minimum were recorded per sample. Sorted cells were recovered in 2X YPD containing 1/100 penicillin-streptomycin for 12h. Then 20 ug/mL zeocine was supplemented and cells were allowed to recover another 24 hours. 15% glycerol stocks of sorting cells were prepared. Recovered cells previous sorting was used for the next sorting. 100 nM and 50 nM PD-L1 were used for first sorting, 25 nM PD-L1 was used for second sorting, 5 nM and 1 nM PD-L1 were used for third sorting. For the fourth sorting, 5 nM and 1 nM PD-L1 were used to assess the increase in binding population percentage.

4.2.5. Affinity measurements on *P.pastoris* via flow cytometry

Affinity measurements on GS115 were applied as described above. 100 uL of OD_{600} = 1/mL cells were pelleted for each condition. The cell pellets were then resuspended in 100 uL of biotinylated VEGF (SinoBiological, Cat# 11066-H27H-B) over a range of concentrations (0.015 nM-100 nM), incubated for 2 hours and 1/200 anti-flag rabbit antibody (SigmaAldrich F7425) was added after 1 hour incubation of the antigen. Cells were pelleted and washed twice with staining buffer. Cells were pelleted in 100 uL staning buffer and incubated with 1/500 anti-rabbit Alexa Fluor 488 conjugated antibody (LifeTech A11008), 1/400 Streptavidin-PE (Pharmingen 554061). After the incubation with secondary reagents, the cells were rinsed with 200 µL of ice-cold staining buffer and then suspended in 150 µL of ice-cold staining buffer for subsequent flow cytometric analysis. 10000 events at minimum were recorded per sample using a BD FACSMelody flow cytometry cell analyzer.

4.3. Results

Founding study

As a very early investigation on dual-specifics, we presented the results of *in silico* investigation of dual-specific antibodies [211]. In this study, we conducted a structural analysis of six different antibodies, including dual-specifics and parental mono-specific template, aiming to gain insights into the determinants of dual-specificity. These dual-specifics exhibited the capability to selectively interact with two different antigens with varying affinities. Our findings revealed that a specific cluster of residues within the Vernier zone region played a crucial role in conferring dual specificity. A limited amount of intramolecular interactions were observed between a particular Vernier zone, referred to as LV4, and the LCDR1 within mono-specific

template, and a notable interactional alteration occurred, leading to closer contacts between the LV4 and LCDR1 loops in the derived dual-specific antibodies. In this study, we showed that previously underestimated Vernier zone regions might help us to modulate antibody specificity in a controlled manner. We concluded that modulation of Vernier zone – CDR interactions might be a new road to understanding antibody specificity and gaining dual-specificity (Figure 4.2). The results from this chapter were published in Proteins: Structure, Function and Bioinformatics as an original research paper in 2019 [211].

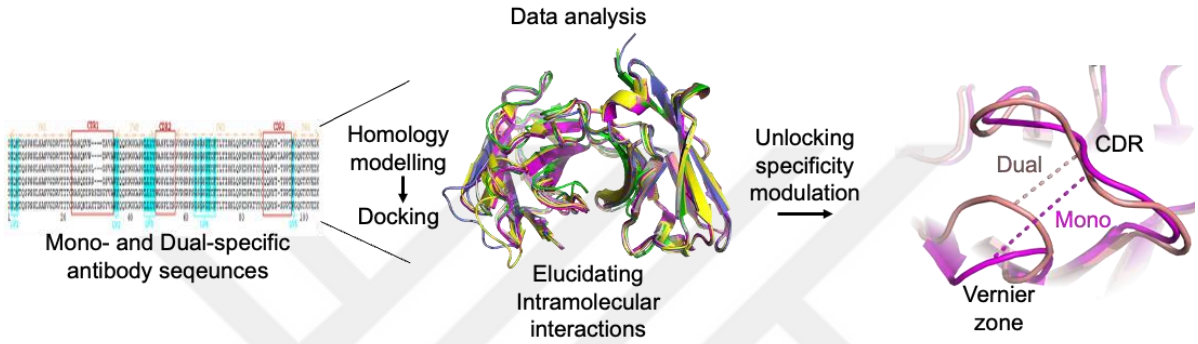


Figure 4. 2 Overview of founding study for dual-specifics

Anti-VEGF scFv displayed via Sag1 system maintains its binding for VEGF

Prior to generation of yeast display libraries and applying sorting procedure, the applicability of the yeast surface display technique via two different cell wall anchor proteins (Sag1 and PpPIR1) for anti-VEGF scFvs was evaluated. The expression and antigen binding activity of the scFv on yeast were tested using an anti-VEGF scFv gene that was generated and characterized at chapter 2 and 3 within the scope of this thesis.

Two different cell wall anchor proteins of the yeast species were used for the surface display of scFv. PpPIR1 display system was orientated as PpPIR1-V_L-linker-V_H-Myc-6xHis **(Figure 4.3A)** that expression could be detected via myc-tag or his-tag. Sag1 display system was orientated as Flag-V_L-linker-V_H-V5-Sag1 **(Figure 4.3B)** that expression could be detected via flag-tag or his-tag. In both system, antigen conjugated biotin and fluophore conjugated streptavidin interaction was utilized to detect antigen binding of the displayed scFvs.

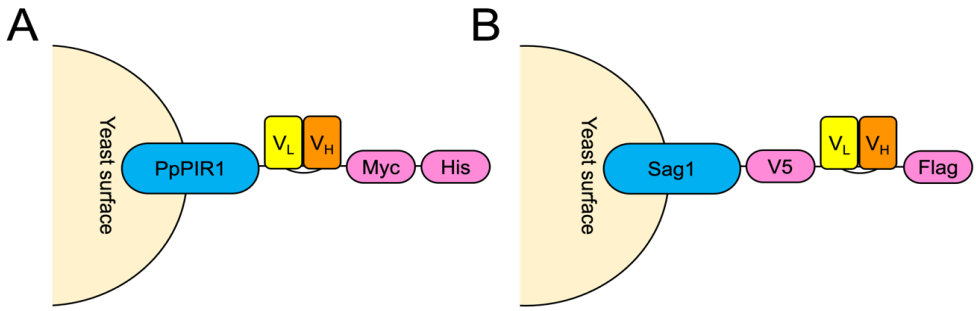


Figure 4. 3 Orientation of the generated yeast surface display plasmids.

A) PpPIR1 display system, B) Sag1 display system.

In PpPIR1 display system, expression was detected on 86 % of the total cell population via his-tag while expression couldn't be detected via myc-tag with two different fluorescent dye **(Figure 4.4A-C)**. Besides, ~ 10 % of the total cell population showed VEGF binding **(Figure 4.4D)**. Since the expression cassette was typically integrated in the genome of *P.pastoris* to obtain stable expression strains , VEGF binding was expected as much as detected expression level.

In Sag1 display system, expression and antigen binding were detected through flag-tag and fluophore conjugate streptavidin. ScFv display level on yeast cells was typically observed in a range ~80-100% percentage of the total population depending on their antibody expression level. Antigen binding ability of the displayed scFvs via Sag1 were also observed similar percentage to detected expression level **(Figure 4.5A-B).**

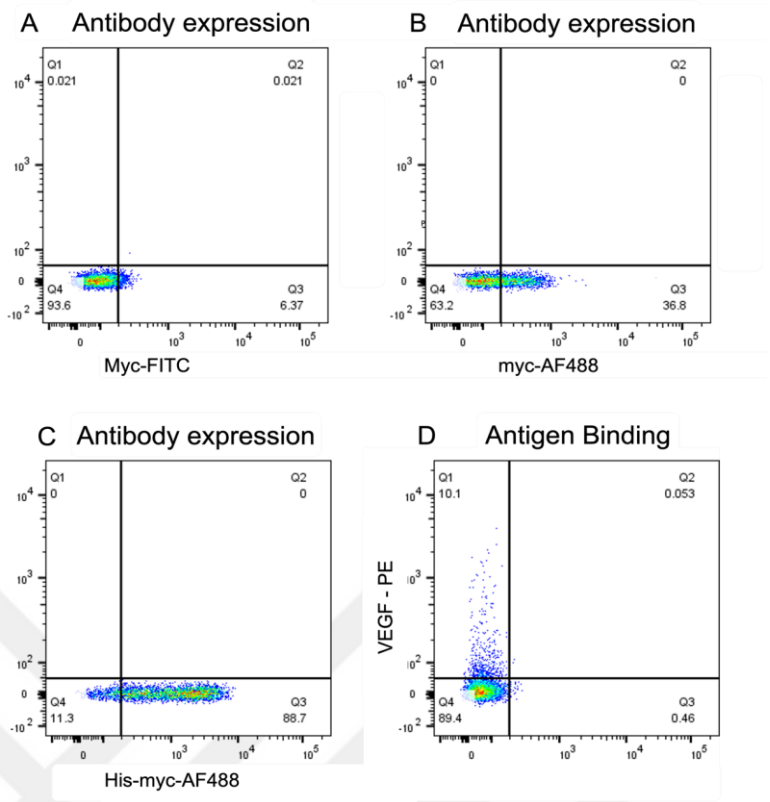


Figure 4. 4 Flow cytometry analysis of scFv display via PpPIR1 cell wall anchor protein.

A) Display level detected through FITC conjugated anti myc-tag antibody, 6.37% of the cell population. B) Display level detected through mouse anti myc-tag antibody, AF488 conjugated anti-mouse antibody, 36.8% of the cell population. C) Display level detected through mouse anti flag-tag antibody, AF488 conjugated anti-mouse antibody, 88.7% of the cell population D) Antigen binding of anti-VEGF scFvs displayed on yeast, 10.1% of the cell population. Biotinylated VEGF and streptavidin PE couple was used to monitor antigen binding. The yeast cells were induced for two days for scFv expression.

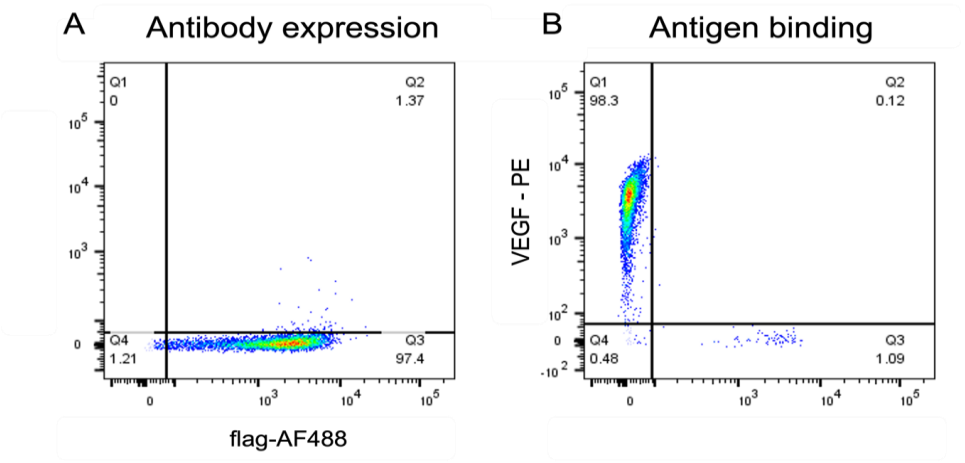


Figure 4. 5 Flow cytometry analysis of scFv display via Sag1 cell wall anchor protein.

A) Display level detected through rabbit anti flag-tag antibody, AF488 conjugated anti rabbit antibody, 97.4 % of the cell population. B) Antigen binding of anti-VEGF scFvs displayed on yeast, 98.7% of the cell population. Biotinylated VEGF and streptavidin PE couple was used to monitor antigen binding. The yeast cells were induced for two days for scFv expression.

As a result, anti-VEGF scFvs were sufficiently displayed on yeast via Sag1 display system and maintained their affinities to VEGF. Thus, all the further studies were carried on utilizing Sag1 display system.

After confirming the utility of Sag1 display system for anti VEGF scFvs, wild type scFv was tested in terms of its antigen binding affinity on the yeast. Binding affinity of soluble scFv was previously analyzed based on it's binding to ligand, VEGF via Surface Plasmon Resonance (SPR) (Chapter 3, Figure 3.6). Binding affinity of surface displayed scFv was analyzed based on its binding to in a range of 0 nM to 100 nM VEGF via flow cytometry at 24h and 48h of induction **(Figure 4.6A-B)**. No substantial difference was observed between the 24-hour and 48-hour induction periods. Binding affinities were determined as 1-5 nM for surface displayed scFv.

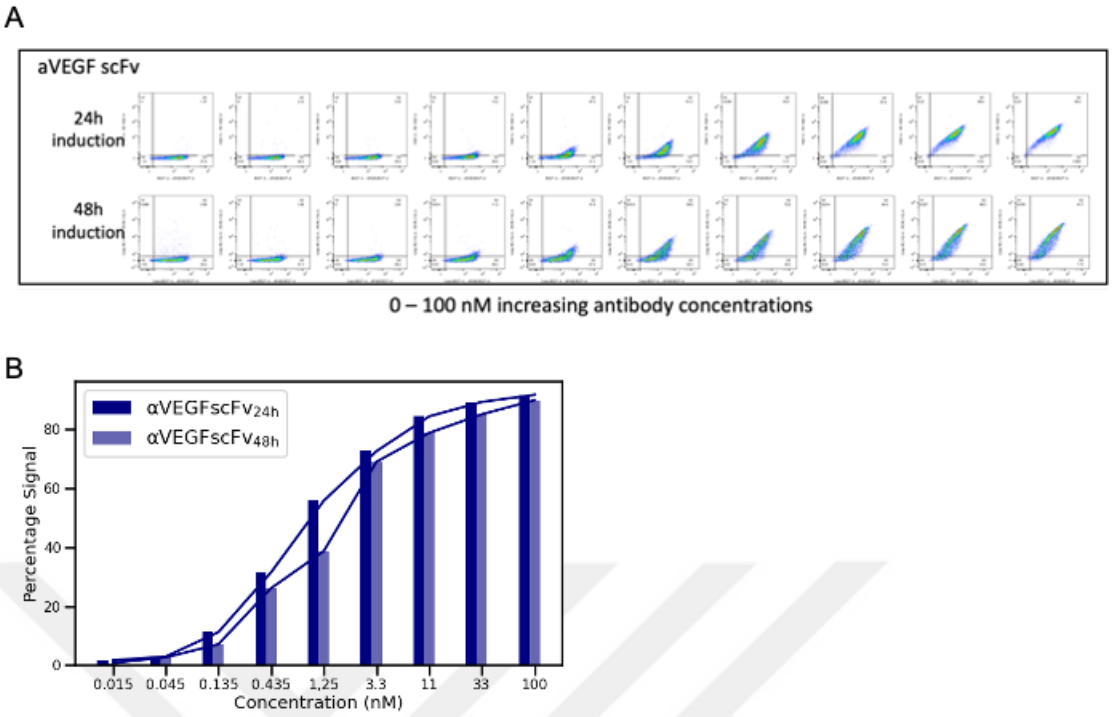


Figure 4. 6 Antigen binding curve of displayed scFv.

A) Flow cytometry graphs of the percentage of antigen binding cell population increases with increasing antigen concentration from 0 nM to 100nM at 24h and 48h induction. B) Bar plot of the percentage of the antigen binding cell population. The expression of scFv was monitored using anti-rabbit anti-flag primary antibody and a goat anti-rabbit secondary antibody (AF488) conjugate. Data were collected from this experiment. Biotinylated VEGF and streptavidin PE couple was used to monitor antigen binding.

Rational behind in silico library design for specificity modulation

In silico analysis was performed to determine paratopes of anti VEGF scFv to its antigen VEGF. Four of the CDRs, HCDR1-3 and LCDR3 of this template showed direct contacts with VEGF (paratope) while LCDR1, LCDR2 and LV4 loops were conformationally distant from the VEGF epitopes and showed no contact with VEGF **(Figure 4.7A-B)**. Thus, a significant amount of light chain paratope was determined available for a second antigen-binding without disturbing the VEGF binding.

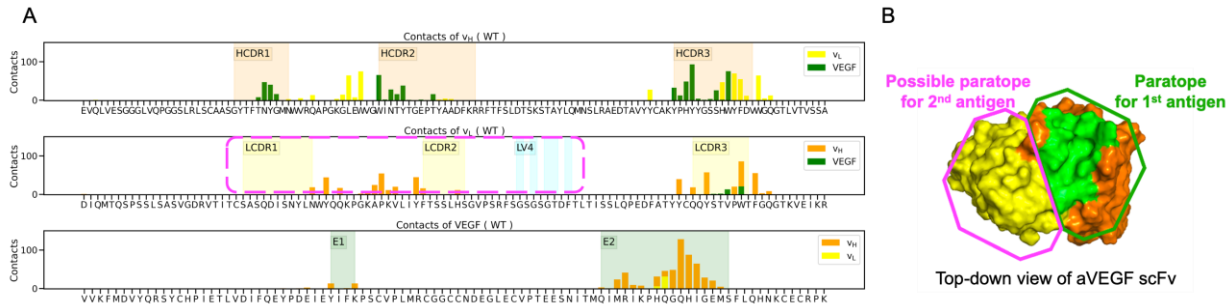


Figure 4. 7 Light paratope is available for second antigen binding.

Interaction plot between anti VEGF scFv and VEGF. A) Heavy chain residues interactions (top panel), Light chain residues interactions (middle panel), VEGF residues interactions (bottom panel). The available paratope is squared in magenta. B) Top-down view of anti-VEGF scFv. VEGF paratope is squared in green, VEGF binding residues are highlighted in green in the scFv structure.

The residues G68 and T69 were identified as critical Vernier zone residues on LV4, responsible for establishing contacts with the neighboring LCDR1. The backbone and side chains of residues 68 and 69 on the LV4 loop were found to be the key components involved in interacting with specific residues in LCDR1. **(Figure 4.8A-B)**. A hypothesis was put forward suggesting that modifying these residues could potentially bring LCDR1 and LV4 closer together, thereby influencing antibody specificity. Based on these, a library was designed by replacing residues 68 and 69 with 20 different amino acids to determine the importance of the residues on antibody specificity modulation. Thus, the total library size was determined as $4x10^2$.

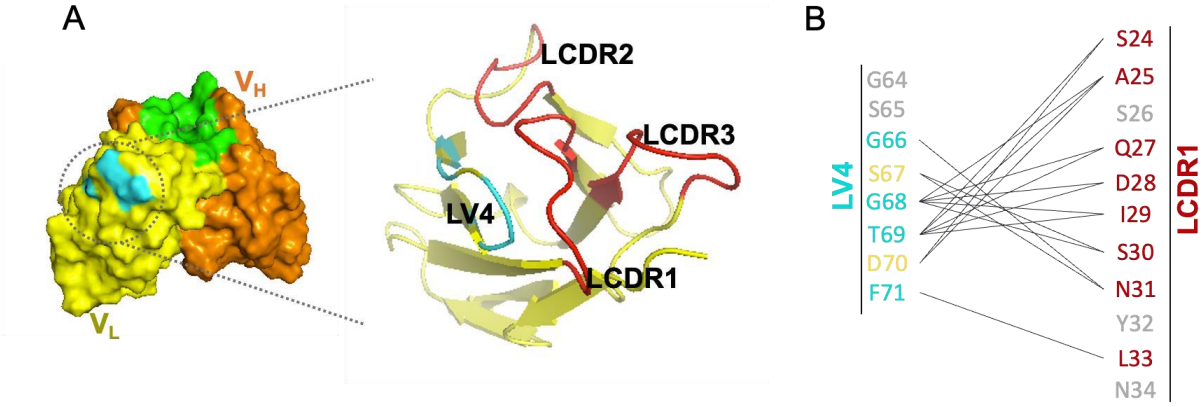


Figure 4. 8 LV4 is the key region for second paratope

A) The structural view of the LCDRs and LV4 region. B) Interaction between LV4 and LCDR1 residues. LV4 faces to LCDR1, interactions are grouped together on positions G68, T69 of LV4. The non-interacting residues are colored in gray.

PD-L1 binding sorting shows enrichment in the library

Prior to sorting of the library, generated library was tested whether it maintains its binding to VEGF. ~ 80 % of the total cell population showed VEGF binding at 10 nM VEGF concentration. It was ensured that designed mutations for specificity modulation did not affect the VEGF binding of the scFv **(Figure 4.9A)**.

Then, a series of sorting steps were conducted to obtain PD-L1 binder scFvs. First, PD-L1 binding flow cytometry analysis was performed at increasing PD-L1 concentrations (5 to 100 nM) to determine starting concentration of the PD-L1 binding sorting in FACS. It was shown that the percentage of the PD-L1 binding cell population was increased in increased antigen concentrations **(Figure 4.9B)**. Although highest percentage of the cell population was observed at 100 nM PD-L1 incubation, to be able to exclude the non-specific binders, 50 nM of PD-L1 concentration was chosen as starting concentration and sorted cells were carried forward in the selection process.

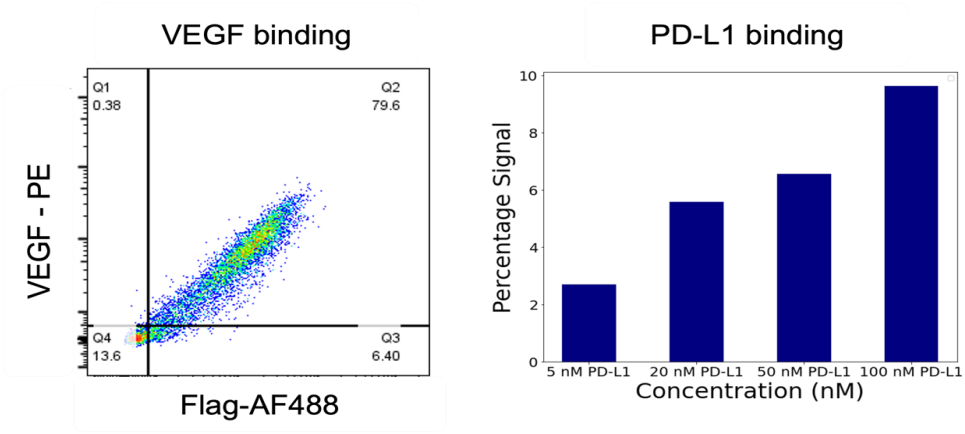


Figure 4. 9 VEGF and PD-L1 binding of generated library

A) Generated library maintained its binding to VEGF. B) The bar plot of the percentage of the PD-L1 binding scFv population in the library.

Further to enrich the PD-L1 binding scFv variants in the libraries, four affinity sorts were performed using FACS (Sort 1 to Sort 4). In each sorting, enriched libraries incubated with decreasing PD-L1 concentrations in a range of 50 nM at Sort 1 (S1), 25 nM at Sort 2 (S2), 1nM at Sort 3 (S3), 1 nM at Sort 4 (S4) respectively. In each sorting, cell population with highest antigen binding signal and highest expression signal were sorted. The selected sorted cell window and the percentage of the sorted cell population were indicated in **Figure 4.10**. While broader cell population were sorted at first two round of the sorting (S1 and S2) in terms of screening all the possible PD-L1 binding candidates, it was shifted to narrowed sorting window at S3 to be able to distinguish the binders. In the final round of PD-L1 affinity sorting, S3 and S4 libraries were compared to each other in terms of their PD-L1 binding affinity at same concentration (1 nM) to observe that the S3 library is an enriched distinct population that possesses a higher binding affinity **(Figure 4.10)**. As a result, sorted libraries at decreasing PD-L1 concentrations in each sorting were enriched a distinct population that showed a better binding affinity than the one before. After the S3 and S4 ibrary recoveries, cells were spreaded on agar plates and a number of clones were sent for sequencing.

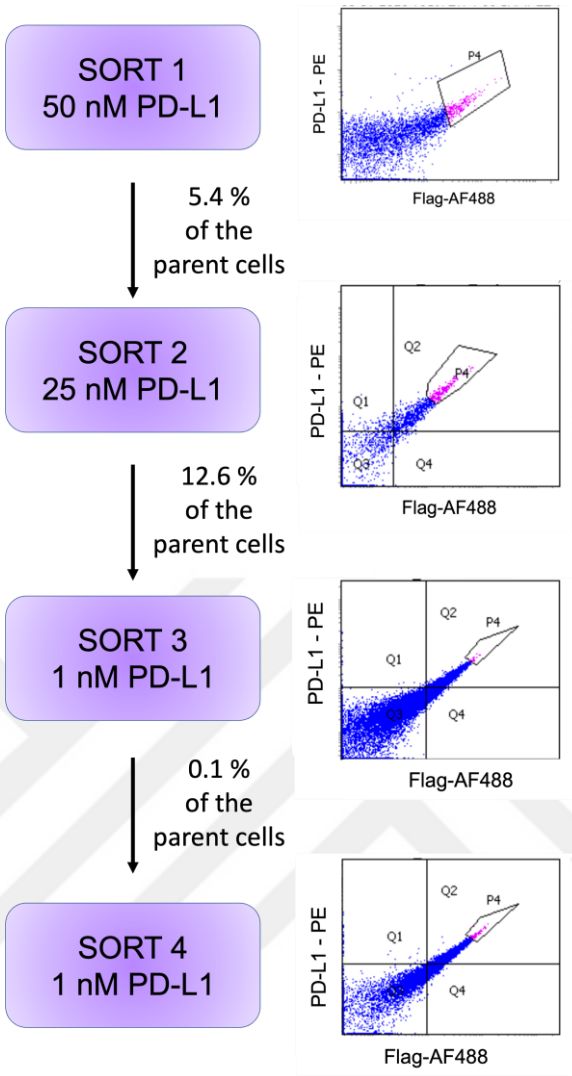
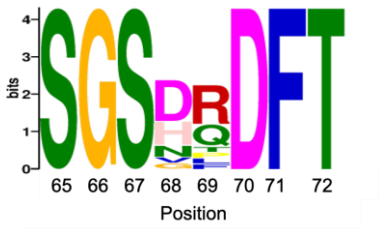


Figure 4. 10 A distinct PD-L1 binding population was enriched in the library.

Mouse anti flag-tag primary antibody and anti-mouse AF 488 conjugated antibody were used to monitor scFv expression. Biotinylated VEGF and streptavidin PE couple was used to monitor VEGF binding.

Unique variants from S4 libraries were taken into further investigation. The unique motifs that obtained at least two times in the population, counted for enrichment. The clones that were enriched but had point mutations on other positions did not include. The most enriched motifs were determined as DR, 40% of the population followed by HQ and NQ, 13 % population at those positions (**Figure 4.11A-B**).

A



B

Position	V_L-68	V_L-69	Presence
C1	D	R	40.0 %
C2	H	Q	13.3 %
C3	N	Q	13.3 %

Figure 4. 11 Enriched motifs from the synthetic library.

A) Glam2 analysis of the clones. B) The presence of the most enriched motifs, DR, HQ and NQ.

Further these clones, DR, HQ and NQ, were analyzed via flow cytometry whether if they showed binding difference to VEGF and PD-L1 compared to WT. The fluorescence signal percentage of the VEGF binding (at 10 nM) of enriched clones and WT were similar to each other that showed no significance in ANOVA test (**Figure 4.12A**). On the other hand, the clone with NQ motif showed significant PD-L1 binding (at 100 nM) compared to WT. The other enriched clones DR and HQ were resulted same PD-L1 binding to WT (**Figure 4.12B**). Here it was shown that VEGF binding of clones preserved with those mutations, while PD-L1 binding was gained with NQ mutations.

Subsequently, flow cytometric analysis was employed to assess the antigen binding affinity of the NQ clone (with altered residues) and the WT clone towards VEGF and PD-L1. The binding affinities of clone NQ and WT to VEGF were found to be very similar, with the highest value of 4.40 ± 0.90 for NQ and 5.44 ± 0.90 for WT. On other hand, the binding affinity of clone NQ to PD-L1 was determined 51.21 ± 14.03 while WT binding to PD-L1 could not be fitted due to having flat-line data in increasing concentration of the PD-L1 **(Figure 4.13A-B).**

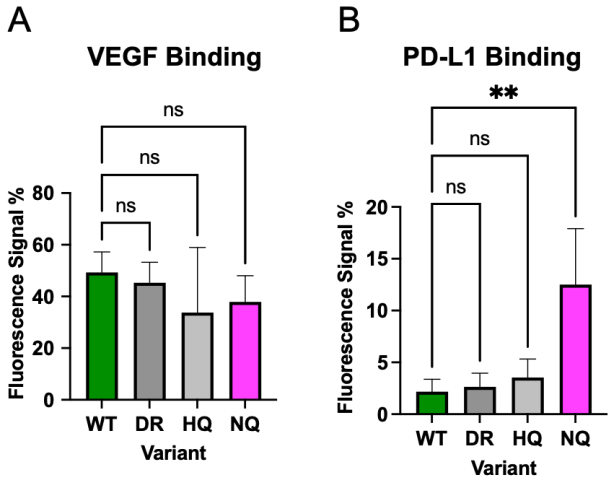


Figure 4. 12 The VEGF and PD-L1 binding of enriched clones and WT.

The VEGF (A) and PD-L1 (B) binding of enriched clones and WT. The ANOVA test showed that the PD-L1 binding of the clone with NQ motif found to be significant compared to WT while VEGF binding of all clones and WT remained similar, that is statistically non-significant.

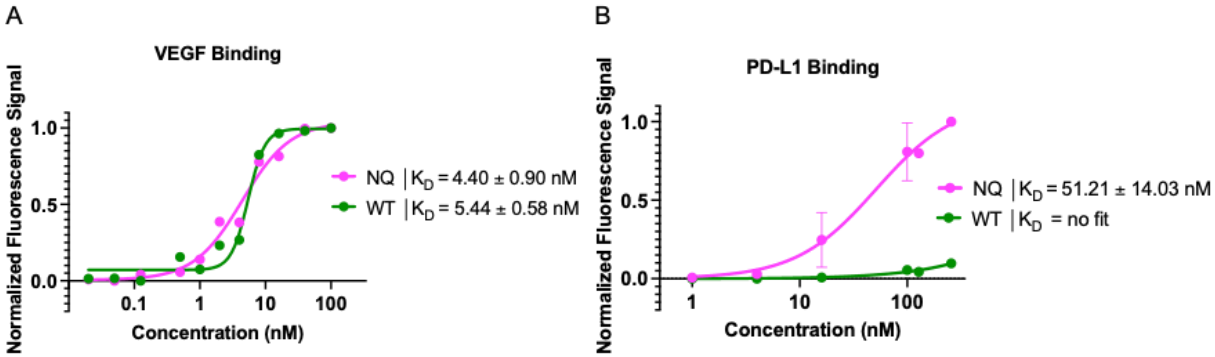


Figure 4. 13 Binding affinities of the clone NQ and WT for VEGF and PD-L1.

4.4. Discussion

In this chapter, we aimed to evaluate the role of an antigen facing loop, Light Vernier 4 (LV4) in the modulation of antibody specificity through one of the directed evolution strategy, yeast

display system and develop a dual-specific scFv antibody fragment. For this purpose, a monospecific anti VEGF scFv fragment that has been designed from a full-length monoclonal antibody, produced, and characterized with high stability and affinity, detailed in chapter 3, was used as a monospecific template. Although four of this template's CDRs are involved in antigen binding, the LCDR1 and LCDR2 loops are conformationally distant from VEGF epitopes and have no contact with VEGF. Thus, a significant amount of light chain paratope is available for a second antigen binding without disturbing VEGF binding.

"Programmed-death-ligand-1, (PD-L1)" is overexpressed in cancer cells suppressing the T-cell-mediated anti-tumor response, thus blocking PD-L1 becomes important in immune checkpoint blocking strategies. Moreover, VEGF signaling plays a vital role in the immunosuppressive tumor microenvironment as well as its function in tumor angiogenesis. Combination therapy approaches have great potential to target both PD-L1 and VEGF. Thus, VEGF/PD-L1 dual targeting is chosen as a case study.

Generating protein variants with single mutations and/or their combinations one by one and testing the effects of these mutations on desired features is a laborious, time-consuming, and costly process. To address this, we generated a synthetic library by utilizing *in silico* analysis and rational design approaches. On the other hand, preparing a purpose based synthetic library is not enough to be able to obtain antibody fragments with desired characteristics. Besides that, it is essential to employ a screening step to be able to sort those variants with the desired characteristics from the generated libraries. Yeast surface display is considered one of the most effective screening techniques for such purposes due to its ability to facilitate rapid screening of antibody libraries, allowing for the identification of antibodies with desired characteristics.

Here we first evaluated the applicability of the yeast surface display technique via two different cell wall anchor proteins, PpPIR1 and Sag1, for anti VEGF scFv. We verified the scFv expression level and antigen binding activity of the scFvs on the yeast via Sag1. The scFv display level on yeast cells was typically in a range of ~90-100. In addition to that, the findings showed that the anti VEGF scFvs maintain their binding affinities to VEGF. Following validation of the yeast surface display system for anti VEGF scFvs, antigen binding of wild type scFv were assessed on yeast. The results showed that surface displayed scFv binds to VEGF at 24 hours and 48 hours of induction similar, shows similar binding properties. Next, we

evaluated the starting concentration of the PD-L1 for sorting process. For that, we performed a flow cytometric analysis to assess the binding activity of scFv library against PD-L1 at increasing concentrations, 5 nM to 100 nM. The results showed that the PD-L1 binding signal increased as the PD-L1 concentration increased. However, as we were aiming to sort those scFv antibodies from the libraries that specific to PD-L1, we concluded that using 50 nM PD-L1 would be efficient to cover all possible specific binding candidates while excluding the non-specific binders. We also evaluated the library whether it maintained its binding to VEGF. It was observed that scFv library maintained its binding to VEGF after the possible two mutations on the LV4 loop.

A series of sorting steps were conducted to obtain VEGF-PD-L1 dual binding scFvs. After first two round of affinity sorting at high PD-L1 concentrations (50 nM and 25 nM), Last two rounds of FACS with low concentration PD-L1 (1nM) was applied to only enrich the PD-L1 specific variants in the libraries. During the last two rounds of the FACS, Sort 3 and Sort 4, top 1% population (~1-10x10^3 cells) were sorted that showed highest expression and PD-L1 binding signal were sorted. The recovered Sort 4 library were further plated. Then, a number of unique variants from Sort 4 library were taken into investigation in terms of sequence determination by sanger sequencing and *in silico* analysis. Sequencing resulted that we enriched 3 motifs in the library DR, HQ and NQ. When we compared the VEGF binding of these clones with WT, the clones maintained their VEGF binding after the mutations on light chain. The difference between VEGF binding of WT and the clones was found to be insignificant by ANOVA test. Upon comparing the PD-L1 binding of these clones with the WT clone, it was observed that one of the clones, namely NQ, exhibited significantly higher binding to PD-L1. The PD-L1 binding of the other clones remained as same as WT.

Finally, we verified the VEGF and PD-L1 binding of promising clone NQ by determining the binding kinetics, K_D, in increasing concentration of the antigens. The binding affinities were found very similar to VEGF while only clone NQ showed increasing binding to increasing concentration of PD-L1.

5. Conclusion and Future Perspectives

In this thesis, first, we generated an scFv fragment targeting VEGF improved the developability properties, affinity and stability, through altering non-CDR residues. Further we focused on the antibody specificity modulation approaches. We investigated the role of a non-CDR region,

Vernier zone, on antibody specificity. Second, we established a novel antibody engineering strategy to gain dual-specificity through modulation of non-CDR regions, and as a result we generate a dual-specific scFv fragment against VEGF and PD-L1 for dual blockade of angiogenesis and immune checkpoint for advanced cancer therapies.

The importance of the dual binding strategies is increased due to their potential. For future studies, this study gives insights on the concept of modulation of mono-specificity to dual-specificity for any antibody.

REFERENCES

1. C.A. Janeway Jr., T., M. Walport, et al.,, The structure of a typical antibody molecule, in Immunobiology: The Immune System in Health and Disease. 2001, Garland Science: New York.
2. Smith, K., et al., Antigen nature and complexity influence human antibody light chain usage and specificity. Vaccine, 2016. **34**(25): p. 2813-20.
3. Vessiere-Louveaux, F.M., W. Hijmans, and H.R. Schuit, Multiple heavy chain isotypes on the membrane of the small B lymphocytes in human blood. Clin Exp Immunol, 1981. **43**(1): p. 149-56.
4. Poljak, R.J., et al., Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2,8-A resolution. Proc Natl Acad Sci U S A, 1973. **70**(12): p. 3305-10.
5. Porter, R.R., The hydrolysis of rabbit y-globulin and antibodies with crystalline papain. Biochem J, 1959. **73**(1): p. 119-26.
6. Kiyoshi, M., et al., Glycosylation of IgG-Fc: a molecular perspective. Int Immunol, 2017. **29**(7): p. 311-317.
7. Teplyakov, A., et al., IgG2 Fc structure and the dynamic features of the IgG CH2-CH3 interface. Mol Immunol, 2013. **56**(1-2): p. 131-9.
8. Saphire, E.O., et al., Contrasting IgG structures reveal extreme asymmetry and flexibility. J Mol Biol, 2002. **319**(1): p. 9-18.
9. Lesk, A.M. and C. Chothia, Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint. Nature, 1988. **335**(6186): p. 188-90.
10. Berek, C. and C. Milstein, Mutation drift and repertoire shift in the maturation of the immune response. Immunol Rev, 1987. **96**: p. 23-41.
11. Chothia, C. and A.M. Lesk, Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol, 1987. **196**(4): p. 901-17.
12. Atassi, M.Z., Antigenic structures of proteins. Their determination has revealed important aspects of immune recognition and generated strategies for synthetic mimicking of protein binding sites. Eur J Biochem, 1984. **145**(1): p. 1-20.
13. Karadag, M., et al., Physicochemical determinants of antibody-protein interactions. Adv Protein Chem Struct Biol, 2020. **121**: p. 85-114.
14. Kabat, E.A. and T.T. Wu, Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. J Immunol, 1991. **147**(5): p. 1709-19.
15. Ewert, S., A. Honegger, and A. Pluckthun, Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. Methods, 2004. **34**(2): p. 184-99.
16. Honegger, A., et al., The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains. Protein Eng Des Sel, 2009. **22**(3): p. 121-34.
17. Presta, L.G., et al., Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders. Cancer Res, 1997. **57**(20): p. 4593-9.
18. Yu, Y., et al., A humanized anti-VEGF rabbit monoclonal antibody inhibits angiogenesis and blocks tumor growth in xenograft models. PLoS One, 2010. **5**(2): p. e9072.
19. Bogen, J.P., et al., Humanization of Chicken-Derived Antibodies by Yeast Surface Display. Methods Mol Biol, 2022. **2491**: p. 335-360.

20. Jones, P.T., et al., Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature, 1986. **321**(6069): p. 522-5.
21. Pelat, T., et al., Germline humanization of a non-human primate antibody that neutralizes the anthrax toxin, by in vitro and in silico engineering. J Mol Biol, 2008. **384**(5): p. 1400-7.
22. Queen, C., et al., A humanized antibody that binds to the interleukin 2 receptor. Proc Natl Acad Sci U S A, 1989. **86**(24): p. 10029-33.
23. Pavlinkova, G., et al., Effects of humanization and gene shuffling on immunogenicity and antigen binding of anti-TAG-72 single-chain Fvs. Int J Cancer, 2001. **94**(5): p. 717-26.
24. Kugler, M., et al., Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework. Protein Eng Des Sel, 2009. **22**(3): p. 135-47.
25. Riley, C.J., et al., Design and activity of a murine and humanized anti-CEACAM6 single-chain variable fragment in the treatment of pancreatic cancer. Cancer Res, 2009. **69**(5): p. 1933- 40.
26. Riechmann, L., et al., Reshaping human antibodies for therapy. Nature, 1988. **332**(6162): p. 323-7.
27. Foote, J. and G. Winter, Antibody framework residues affecting the conformation of the hypervariable loops. J Mol Biol, 1992. **224**(2): p. 487-99.
28. Safdari, Y., et al., Antibody humanization methods - a review and update. Biotechnol Genet Eng Rev, 2013. **29**: p. 175-86.
29. Makabe, K., et al., Thermodynamic consequences of mutations in vernier zone residues of a humanized anti-human epidermal growth factor receptor murine antibody, 528. J Biol Chem, 2008. **283**(2): p. 1156-66.
30. Schroeder, H.W., Jr., Similarity and divergence in the development and expression of the mouse and human antibody repertoires. Dev Comp Immunol, 2006. **30**(1-2): p. 119-35.
31. Chothia, C., et al., Conformations of immunoglobulin hypervariable regions. Nature, 1989. **342**(6252): p. 877-83.
32. MacCallum, R.M., A.C. Martin, and J.M. Thornton, Antibody-antigen interactions: contact analysis and binding site topography. J Mol Biol, 1996. **262**(5): p. 732-45.
33. Ramaraj, T., et al., Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. Biochim Biophys Acta, 2012. **1824**(3): p. 520- 32.
34. Peng, H.P., et al., Origins of specificity and affinity in antibody-protein interactions. Proc Natl Acad Sci U S A, 2014. **111**(26): p. E2656-65.
35. Mian, I.S., A.R. Bradwell, and A.J. Olson, Structure, function and properties of antibody binding sites. J Mol Biol, 1991. **217**(1): p. 133-51.
36. Koide, S. and S.S. Sidhu, The importance of being tyrosine: lessons in molecular recognition from minimalist synthetic binding proteins. ACS Chem Biol, 2009. **4**(5): p. 325-34.
37. Pollard, T.D., A guide to simple and informative binding assays. Mol Biol Cell, 2010. **21**(23): p. 4061-7.
38. Janin, J., A minimal model of protein-protein binding affinities. Protein Sci, 2014. **23**(12): p. 1813-7.
39. Hearty, S., P. Leonard, and R. O'Kennedy, Measuring antibody-antigen binding kinetics using surface plasmon resonance. Methods Mol Biol, 2012. **907**: p. 411-42.
40. Livingstone, J.R., Antibody characterization by isothermal titration calorimetry. Nature, 1996. **384**(6608): p. 491-2.

41. Williamson, M.P., Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc, 2013. **73**: p. 1-16.
42. Yu, X., et al., Development and validation of a cell-based fluorescent method for measuring antibody affinity. J Immunol Methods, 2017. **442**: p. 49-53.
43. Wienken, C.J., et al., Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun, 2010. **1**: p. 100.
44. Abdiche, Y., et al., Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. Anal Biochem, 2008. **377**(2): p. 209-17.
45. Friguet, B., et al., Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. J Immunol Methods, 1985. **77**(2): p. 305- 19.
46. Xue, L.C., et al., PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics, 2016. **32**(23): p. 3676-3678.
47. Myung, Y., et al., mCSM-AB2: guiding rational antibody design using graph-based signatures.
Bioinformatics, 2020. **36**(5): p. 1453-1459.
48. Dehouck, Y., et al., BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations. Nucleic Acids Res, 2013. **41**(Web Server issue): p. W333-9.
49. Zhang, N., et al., MutaBind2: Predicting the Impacts of Single and Multiple Mutations on Protein-Protein Interactions. iScience, 2020. **23**(3): p. 100939.
50. Eisen, H.N., Specificity and degeneracy in antigen recognition: yin and yang in the immune system. Annu Rev Immunol, 2001. **19**: p. 1-21.
51. Cohn, M., Degeneracy, mimicry and crossreactivity in immune recognition. Mol Immunol, 2005. **42**(5): p. 651-5.
52. Bostrom, J., et al., High affinity antigen recognition of the dual specific variants of herceptin is entropy-driven in spite of structural plasticity. PLoS One, 2011. **6**(4): p. e17887.
53. Wardemann, H., et al., Predominant autoantibody production by early human B cell precursors. Science, 2003. **301**(5638): p. 1374-7.
54. Van Regenmortel, M.H., Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. J Mol Recognit, 2014. **27**(11): p. 627-39.
55. Dimitrov, J.D., et al., Antibody polyreactivity in health and disease: statu variabilis. J Immunol,
2013. **191**(3): p. 993-9.
56. Jain, D. and D.M. Salunke, Antibody specificity and promiscuity. Biochem J, 2019. **476**(3): p. 433-447.
57. Cohn, M., A new concept of immune specificity emerges from a consideration of the self-nonself discrimination. Cell Immunol, 1997. **181**(2): p. 103-8.
58. Parnes, O., From interception to incorporation: degeneracy and promiscuous recognition as precursors of a paradigm shift in immunology. Mol Immunol, 2004. **40**(14-15): p. 985-91.
59. Jain, T., et al., Biophysical properties of the clinical-stage antibody landscape. Proc Natl Acad Sci U S A, 2017. **114**(5): p. 944-949.
60. Starr, C.G. and P.M. Tessier, Selecting and engineering monoclonal antibodies with drug-like specificity. Curr Opin Biotechnol, 2019. **60**: p. 119-127.
61. Lee, H.Y., et al., "Two-in-One" approach for bioassay selection for dual specificity antibodies. J Immunol Methods, 2017. **448**: p. 74-79.
62. Kaleli, N.E., M. Karadag, and S. Kalyoncu, Phage display derived therapeutic antibodies have enriched aliphatic content: Insights for developability issues. Proteins, 2019. **87**(7): p. 607-618.
63. Chuang, G.Y., et al., Eliminating antibody polyreactivity through addition of N-linked glycosylation. Protein Sci, 2015. **24**(6): p. 1019-30.

64. Igawa, T., et al., Reduced elimination of IgG antibodies by engineering the variable region. Protein Eng Des Sel, 2010. **23**(5): p. 385-92.
65. Kelly, R.L., et al., Reduction of Nonspecificity Motifs in Synthetic Antibody Libraries. J Mol Biol,
2018. **430**(1): p. 119-130.
66. Rabia, L.A., et al., Net charge of antibody complementarity-determining regions is a key predictor of specificity. Protein Eng Des Sel, 2018. **31**(11): p. 409-418.
67. Nguyen, A.W.M., J.A., Engineering Antibody-Based Therapeutics: Progress and Opportunities, in Protein Engineering: Tools and Applications,, H. Zhao, Editor. 2021, WILEY-VCH GmbH. p. 317-351.
68. Bhatia, K., Bhumika, and A. Das, Combinatorial drug therapy in cancer - New insights. Life Sci,
2020. **258**: p. 118134.
69. Henricks, L.M., et al., The use of combinations of monoclonal antibodies in clinical oncology. Cancer Treat Rev, 2015. **41**(10): p. 859-67.
70. Corraliza-Gorjon, I., et al., New Strategies Using Antibody Combinations to Increase Cancer Treatment Effectiveness. Front Immunol, 2017. **8**: p. 1804.
71. Mandy, W.J., M.M. Rivers, and A. Nisonoff, Recombination of univalent subunits derived from rabbit antibody. J Biol Chem, 1961. **236**: p. 3221-6.
72. Nisonoff, A. and M.M. Rivers, Recombination of a mixture of univalent antibody fragments of different specificity. Arch Biochem Biophys, 1961. **93**: p. 460-2.
73. Ma, J., et al., Bispecific Antibodies: From Research to Clinical Application. Front Immunol, 2021. **12**: p. 626616.
74. Holliger, P., T. Prospero, and G. Winter, "Diabodies": small bivalent and bispecific antibody fragments. Proc Natl Acad Sci U S A, 1993. **90**(14): p. 6444-8.
75. Kwon, N.Y., Y. Kim, and J.O. Lee, Structural diversity and flexibility of diabodies. Methods, 2019. **154**: p. 136-142.
76. Beckmann, R., et al., DutaFabs are engineered therapeutic Fab fragments that can bind two targets simultaneously. Nat Commun, 2021. **12**(1): p. 708.
77. Behring E, a.K., S., Uber das Zustandekommen Der Diphtherie- Immunitat Und der Tetanus- Immunitat Bei Thieren. Dtsch Med Wochenschr, 1890. **49**.
78. Kohler, G. and C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity. Nature, 1975. **256**(5517): p. 495-7.
79. Padlan, E.A., et al., Structure at 4.5 A resolution of a phosphorylcholine-binding fab. Nat New Biol, 1973. **245**(145): p. 165-7.
80. Poljak, R.J., et al., The three-dimensional structure of the fab' fragment of a human myeloma immunoglobulin at 2.0-angstrom resolution. Proc Natl Acad Sci U S A, 1974. **71**(9): p. 3440-4.
81. Hughes-Jones, N.C. and B. Gardner, Reaction between the isolated globular sub-units of the complement component C1q and IgG-complexes. Mol Immunol, 1979. **16**(9): p. 697-701.
82. Ricklin, D., et al., Complement component C3 - The "Swiss Army Knife" of innate immunity and host defense. Immunol Rev, 2016. **274**(1): p. 33-58.
83. Share & Trends Analysis Report By Product (Primary, S., By Type (Monoclonal Antibodies, Polyclonal Antibodies), By Technology, By Source, By Application, By End-use, By Region, And Segment Forecasts., Research Antibodies Market Size, . 2021.
84. Smith, G.P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, 1985. **228**(4705): p. 1315-7.
85. Hanes, J. and A. Pluckthun, In vitro selection and evolution of functional proteins by using ribosome display. Proc Natl Acad Sci U S A, 1997. **94**(10): p. 4937-42.
86. Nagumo, Y., et al., PURE mRNA display for in vitro selection of single-chain antibodies. J Biochem, 2016. **159**(5): p. 519-26.

87. Boder, E.T. and K.D. Wittrup, Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol, 1997. **15**(6): p. 553-7.
88. Georgiou, G., et al., Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. Nat Biotechnol, 1997. **15**(1): p. 29-34.
89. Ho, M., S. Nagata, and I. Pastan, Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. Proc Natl Acad Sci U S A, 2006. **103**(25): p. 9637-42.
90. Lonberg, N., et al., Antigen-specific human antibodies from mice comprising four distinct genetic modifications. Nature, 1994. **368**(6474): p. 856-9.
91. Mendez, M.J., et al., Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice. Nat Genet, 1997. **15**(2): p. 146-56.
92. Traggiai, E., et al., An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. Nat Med, 2004. **10**(8): p. 871-5.
93. Tiller, T., et al., Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. J Immunol Methods, 2008. **329**(1-2): p. 112-24.
94. Adams, J.J., B. Nelson, and S.S. Sidhu, Recombinant genetic libraries and human monoclonal antibodies. Methods Mol Biol, 2014. **1060**: p. 149-70.
95. Arslan, M., D. Karadag, and S. Kalyoncu, Protein engineering approaches for antibody fragments: directed evolution and rational design approaches. Turk J Biol, 2019. **43**(1): p. 1- 12.
96. Ducancel, F. and B.H. Muller, Molecular engineering of antibodies for therapeutic and diagnostic purposes. MAbs, 2012. **4**(4): p. 445-57.
97. Viegas-Barroso J.F., H., M.E., Whelan, M., EURL ECVAM Recommendation on Non-Animal-Derived Antibodies. Publications Office of the European Union, 2020. **EUR 30185 EN**.
98. Cannon, D.A., et al., Experimentally guided computational antibody affinity maturation with de novo docking, modelling and rational design. PLoS Comput Biol, 2019. **15**(5): p. e1006980.
99. Berman, H.M., et al., The Protein Data Bank. Nucleic Acids Res, 2000. **28**(1): p. 235-42.
100. Swindells, M.B., et al., abYsis: Integrated Antibody Sequence and Structure-Management, Analysis, and Prediction. J Mol Biol, 2017. **429**(3): p. 356-364.
101. Vita, R., et al., The immune epitope database (IEDB) 3.0. Nucleic Acids Res, 2015. **43**(Database issue): p. D405-12.
102. Sormanni, P., F.A. Aprile, and M. Vendruscolo, Third generation antibody discovery methods: in silico rational design. Chem Soc Rev, 2018. **47**(24): p. 9137-9157.
103. Jarasch, A., et al., Developability assessment during the selection of novel therapeutic antibodies. J Pharm Sci, 2015. **104**(6): p. 1885-1898.
104. Chiu, M.L., et al., Antibody Structure and Function: The Basis for Engineering Therapeutics. Antibodies (Basel), 2019. **8**(4).
105. Bradbury, A.R., et al., Beyond natural antibodies: the power of in vitro display technologies. Nat Biotechnol, 2011. **29**(3): p. 245-54.
106. Petrus, M.L.C., et al., A microbial expression system for high-level production of scFv HIV-neutralizing antibody fragments in Escherichia coli. Appl Microbiol Biotechnol, 2019. **103**(21- 22): p. 8875-8888.
107. Spadiut, O., et al., Microbials for the production of monoclonal antibodies and antibody fragments. Trends Biotechnol, 2014. **32**(1): p. 54-60.
108. Bates, A. and C.A. Power, David vs. Goliath: The Structure, Function, and Clinical Prospects of Antibody Fragments. Antibodies (Basel), 2019. **8**(2).
109. Li, Z., et al., Influence of molecular size on tissue distribution of antibody fragments. MAbs,

2016. **8**(1): p. 113-9.
110. Chen, M., et al., Envafolimab - first PD-1/PD-L1 antibody to be administered by subcutaneous injection for microsatellite instability-high or deficient mismatch repair advanced solid tumors. Expert Opin Biol Ther, 2022. **22**(10): p. 1227-1232.
111. Tadayoni, R., et al., Brolucizumab: A Newly Developed Anti-VEGF Molecule for the Treatment of Neovascular Age-Related Macular Degeneration. Ophthalmologica, 2021. **244**(2): p. 93- 101.
112. Klasse, P.J., Neutralization of Virus Infectivity by Antibodies: Old Problems in New Perspectives. Adv Biol, 2014. **2014**.
113. Kholodenko, R.V., et al., Antibody Fragments as Potential Biopharmaceuticals for Cancer Therapy: Success and Limitations. Curr Med Chem, 2019. **26**(3): p. 396-426.
114. Huang, L., X. Su, and H.J. Federoff, Single-chain fragment variable passive immunotherapies for neurodegenerative diseases. Int J Mol Sci, 2013. **14**(9): p. 19109-27.
115. Demarest, S.J. and S.M. Glaser, Antibody therapeutics, antibody engineering, and the merits of protein stability. Curr Opin Drug Discov Devel, 2008. **11**(5): p. 675-87.
116. Jager, M. and A. Pluckthun, Domain interactions in antibody Fv and scFv fragments: effects on unfolding kinetics and equilibria. FEBS Lett, 1999. **462**(3): p. 307-12.
117. Lundahl, M.L.E., et al., Aggregation of protein therapeutics enhances their immunogenicity:
causes and mitigation strategies. RSC Chem Biol, 2021. **2**(4): p. 1004-1020.
118. Kontermann, R.E., Strategies to extend plasma half-lives of recombinant antibodies. BioDrugs,
2009. **23**(2): p. 93-109.
119. Kang, T.H. and S.T. Jung, Reprogramming the Constant Region of Immunoglobulin G Subclasses for Enhanced Therapeutic Potency against Cancer. Biomolecules, 2020. **10**(3).
120. Stork, R., et al., Biodistribution of a bispecific single-chain diabody and its half-life extended derivatives. J Biol Chem, 2009. **284**(38): p. 25612-9.
121. Frenzel, A., M. Hust, and T. Schirrmann, Expression of recombinant antibodies. Front Immunol, 2013. **4**: p. 217.
122. Rosano, G.L. and E.A. Ceccarelli, Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol, 2014. **5**: p. 172.
123. Bentley, W.E., et al., Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnol Bioeng, 1990. **35**(7): p. 668-81.
124. Sezonov, G., D. Joseleau-Petit, and R. D'Ari, Escherichia coli physiology in Luria-Bertani broth. J Bacteriol, 2007. **189**(23): p. 8746-9.
125. Freudl, R., Signal peptides for recombinant protein secretion in bacterial expression systems. Microb Cell Fact, 2018. **17**(1): p. 52.
126. Baneyx, F., Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol, 1999. **10**(5): p. 411-21.
127. Skerra, A. and A. Pluckthun, Assembly of a functional immunoglobulin Fv fragment in Escherichia coli. Science, 1988. **240**(4855): p. 1038-41.
128. Chen, L.H., et al., Expression, purification, and in vitro refolding of a humanized single-chain Fv antibody against human CTLA4 (CD152). Protein Expr Purif, 2006. **46**(2): p. 495-502.
129. Reddy Chichili, V.P., V. Kumar, and J. Sivaraman, Linkers in the structural biology of protein-protein interactions. Protein Sci, 2013. **22**(2): p. 153-67.
130. Thompson, J., et al., Improved binding of a bivalent single-chain immunotoxin results in increased efficacy for in vivo T-cell depletion. Protein Eng, 2001. **14**(12): p. 1035-41.
131. Yusakul, G., et al., Effect of linker length between variable domains of single chain variable fragment antibody against daidzin on its reactivity. Biosci Biotechnol Biochem, 2016. **80**(7): p.

1306-12.
132. Chen, X., J.L. Zaro, and W.C. Shen, Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev, 2013. **65**(10): p. 1357-69.
133. Volkel, T., et al., Optimized linker sequences for the expression of monomeric and dimeric bispecific single-chain diabodies. Protein Eng, 2001. **14**(10): p. 815-23.
134. Zhang, K., et al., Comprehensive optimization of a single-chain variable domain antibody fragment as a targeting ligand for a cytotoxic nanoparticle. MAbs, 2015. **7**(1): p. 42-52.
135. Klein, J.S., et al., Design and characterization of structured protein linkers with differing flexibilities. Protein Eng Des Sel, 2014. **27**(10): p. 325-30.
136. Navabi, P., et al., Designing and generating a single-chain fragment variable (scFv) antibody against IL2Ralpha (CD25): An in silico and in vitro study. Iran J Basic Med Sci, 2021. **24**(3): p. 360-368.
137. Trinh, R., et al., Optimization of codon pair use within the (GGGGS)3 linker sequence results in enhanced protein expression. Mol Immunol, 2004. **40**(10): p. 717-22.
138. Long, N.E., et al., Linker engineering in anti-TAG-72 antibody fragments optimizes biophysical properties, serum half-life, and high-specificity tumor imaging. J Biol Chem, 2018. **293**(23): p. 9030-9040.
139. Hudson, P.J. and A.A. Kortt, High avidity scFv multimers; diabodies and triabodies. J Immunol Methods, 1999. **231**(1-2): p. 177-89.
140. Le Gall, F., et al., Di-, tri- and tetrameric single chain Fv antibody fragments against human CD19: effect of valency on cell binding. FEBS Lett, 1999. **453**(1-2): p. 164-8.
141. Hennecke, F., C. Krebber, and A. Pluckthun, Non-repetitive single-chain Fv linkers selected by selectively infective phage (SIP) technology. Protein Eng, 1998. **11**(5): p. 405-10.
142. Wen, D., et al., Discovery and investigation of O-xylosylation in engineered proteins containing a (GGGGS)n linker. Anal Chem, 2013. **85**(9): p. 4805-12.
143. Wilting, J. and B. Christ, Embryonic angiogenesis: a review. Naturwissenschaften, 1996. **83**(4): p. 153-64.
144. DiPietro, L.A., Angiogenesis and wound repair: when enough is enough. J Leukoc Biol, 2016. **100**(5): p. 979-984.
145. Reynolds, L.P., A.T. Grazul-Bilska, and D.A. Redmer, Angiogenesis in the female reproductive organs: pathological implications. Int J Exp Pathol, 2002. **83**(4): p. 151-63.
146. Ji, Y., et al., Signaling pathways in the development of infantile hemangioma. J Hematol Oncol, 2014. **7**: p. 13.
147. Koch, A.E. and O. Distler, Vasculopathy and disordered angiogenesis in selected rheumatic diseases: rheumatoid arthritis and systemic sclerosis. Arthritis Res Ther, 2007. **9 Suppl 2**(Suppl 2): p. S3.
148. Fadini, G.P., et al., Angiogenic Abnormalities in Diabetes Mellitus: Mechanistic and Clinical Aspects. J Clin Endocrinol Metab, 2019. **104**(11): p. 5431-5444.
149. Zhang, A., et al., Role of VEGF-A and LRG1 in Abnormal Angiogenesis Associated With Diabetic Nephropathy. Front Physiol, 2020. **11**: p. 1064.
150. Hanahan, D. and R.A. Weinberg, Hallmarks of cancer: the next generation. Cell, 2011. **144**(5): p. 646-74.
151. Ribatti, D. and F. Pezzella, Overview on the Different Patterns of Tumor Vascularization. Cells, 2021. **10**(3).
152. Leung, D.W., et al., Vascular endothelial growth factor is a secreted angiogenic mitogen. Science, 1989. **246**(4935): p. 1306-9.
153. Gupta, M.K. and R.Y. Qin, Mechanism and its regulation of tumor-induced angiogenesis. World J Gastroenterol, 2003. **9**(6): p. 1144-55.
154. Yang, J., J. Yan, and B. Liu, Targeting VEGF/VEGFR to Modulate Antitumor Immunity. Front Immunol, 2018. **9**: p. 978.

155. Inai, T., et al., Inhibition of vascular endothelial growth factor (VEGF) signaling in cancer causes loss of endothelial fenestrations, regression of tumor vessels, and appearance of basement membrane ghosts. Am J Pathol, 2004. **165**(1): p. 35-52.
156. Kong, D.H., et al., A Review of Anti-Angiogenic Targets for Monoclonal Antibody Cancer Therapy. Int J Mol Sci, 2017. **18**(8).
157. Zirlik, K. and J. Duyster, Anti-Angiogenics: Current Situation and Future Perspectives. Oncol Res Treat, 2018. **41**(4): p. 166-171.
158. Antibody Society, T.m.a.a.o.i.r.r., date accessed: 21 April 2023.
159. Ferrara, N., et al., Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. Nat Rev Drug Discov, 2004. **3**(5): p. 391-400.
160. Muller, Y.A., et al., VEGF and the Fab fragment of a humanized neutralizing antibody: crystal structure of the complex at 2.4 A resolution and mutational analysis of the interface. Structure, 1998. **6**(9): p. 1153-67.
161. Brozzo, M.S., et al., Thermodynamic and structural description of allosterically regulated VEGFR-2 dimerization. Blood, 2012. **119**(7): p. 1781-8.
162. Gerber, H.P., et al., Complete inhibition of rhabdomyosarcoma xenograft growth and neovascularization requires blockade of both tumor and host vascular endothelial growth factor. Cancer Res, 2000. **60**(22): p. 6253-8.
163. Warren, R.S., et al., Regulation by vascular endothelial growth factor of human colon cancer tumorigenesis in a mouse model of experimental liver metastasis. J Clin Invest, 1995. **95**(4): p. 1789-97.
164. Kabbinavar, F., et al., Phase II, randomized trial comparing bevacizumab plus fluorouracil (FU)/leucovorin (LV) with FU/LV alone in patients with metastatic colorectal cancer. J Clin Oncol, 2003. **21**(1): p. 60-5.
165. Adding a humanized antibody to vascular endothelial growth factor (Bevacizumab, Avastin) to chemotherapy improves survival in metastatic colorectal cancer. Clin Colorectal Cancer, 2003. **3**(2): p. 85-8.
166. Kindler, H.L., et al., Phase II trial of bevacizumab plus gemcitabine in patients with advanced pancreatic cancer. J Clin Oncol, 2005. **23**(31): p. 8033-40.
167. Ebos, J.M., et al., Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. Cancer Cell, 2009. **15**(3): p. 232-9.
168. Paez-Ribes, M., et al., Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. Cancer Cell, 2009. **15**(3): p. 220-31.
169. Loges, S., T. Schmidt, and P. Carmeliet, Mechanisms of resistance to anti-angiogenic therapy and development of third-generation anti-angiogenic drug candidates. Genes Cancer, 2010.
1(1): p. 12-25.
170. Li, Y.L., H. Zhao, and X.B. Ren, Relationship of VEGF/VEGFR with immune and cancer cells: staggering or forward? Cancer Biol Med, 2016. **13**(2): p. 206-14.
171. Ziogas, A.C., et al., VEGF directly suppresses activation of T cells from ovarian cancer patients and healthy individuals via VEGF receptor Type 2. Int J Cancer, 2012. **130**(4): p. 857-64.
172. Gavalas, N.G., et al., VEGF directly suppresses activation of T cells from ascites secondary to ovarian cancer via VEGF receptor type 2. Br J Cancer, 2012. **107**(11): p. 1869-75.
173. Arslan, M., et al., Effect of non-repetitive linker on in vitro and in vivo properties of an anti- VEGF scFv. Sci Rep, 2022. **12**(1): p. 5449.
174. Wishart, D.S., et al., DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res, 2018. **46**(D1): p. D1074-D1082.
175. Studier, F.W., Protein production by auto-induction in high density shaking cultures. Protein Expr Purif, 2005. **41**(1): p. 207-34.
176. Arslan M., C.-A.G., Kalyoncu S.,, Antibodyengineering approaches on Vernier zone residues to

develop biobetter antibody therapeutics for cancer treatment. Turkish Journal of Biochemistry, 2019. **Proceeding of the 2nd International Ion Channel and Cancer Congress (IonCC2019)**(44).
177. Expasy. https://web.expasy.org/protscale/pscale/Averageflexibility.html.
178. Bhaskaran, R.P.P.K.a.P., P.K.,, Positional flexibilities of amino acid residues in globular proteins. International Journal of Peptide and Protein Research, 1988. **32**(4): p. 241-255.
179. Calis, J.J., et al., Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Comput Biol, 2013. **9**(10): p. e1003266.
180. Abhinandan, K.R. and A.C. Martin, Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. Mol Immunol, 2008. **45**(14): p. 3832-9.
181. Lei, S.P., et al., Characterization of the Erwinia carotovora pelB gene and its product pectate lyase. J Bacteriol, 1987. **169**(9): p. 4379-83.
182. Stemmer, W.P., DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10747-51.
183. Winter, G., et al., Making antibodies by phage display technology. Annu Rev Immunol, 1994. **12**: p. 433-55.
184. Ongaro, T., et al., A novel format for recombinant antibody-interleukin-2 fusion proteins exhibits superior tumor-targeting properties in vivo. Oncotarget, 2020. **11**(41): p. 3698-3711.
185. Klement, M., et al., Effect of linker flexibility and length on the functionality of a cytotoxic engineered antibody fragment. J Biotechnol, 2015. **199**: p. 90-7.
186. Vihinen, M., E. Torkkila, and P. Riikonen, Accuracy of protein flexibility predictions. Proteins, 1994. **19**(2): p. 141-9.
187. Bigman, L.S. and Y. Levy, Proteins: molecules defined by their trade-offs. Curr Opin Struct Biol, 2020. **60**: p. 50-56.
188. Rabia, L.A., et al., Understanding and overcoming trade-offs between antibody affinity, specificity, stability and solubility. Biochem Eng J, 2018. **137**: p. 365-374.
189. Houlihan, G., et al., Directed evolution of anti-HER2 DARPins by SNAP display reveals stability/function trade-offs in the selection process. Protein Eng Des Sel, 2015. **28**(9): p. 269- 79.
190. Faber, M.S. and T.A. Whitehead, Data-driven engineering of protein therapeutics. Curr Opin Biotechnol, 2019. **60**: p. 104-110.
191. Cohen-Khait, R., et al., Promiscuous Protein Binding as a Function of Protein Stability. Structure, 2017. **25**(12): p. 1867-1874 e3.
192. Mandrup, O.A., et al., A novel heavy domain antibody library with functionally optimized complementarity determining regions. PLoS One, 2013. **8**(10): p. e76834.
193. Xu, J.L. and M.M. Davis, Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. Immunity, 2000. **13**(1): p. 37-45.
194. Bostrom, J., et al., Improving antibody binding affinity and specificity for therapeutic development. Methods Mol Biol, 2009. **525**: p. 353-76, xiii.
195. Herold, E.M., et al., Determinants of the assembly and function of antibody variable domains. Sci Rep, 2017. **7**(1): p. 12276.
196. Zhao, J., R. Nussinov, and B. Ma, The Allosteric Effect in Antibody-Antigen Recognition. Methods Mol Biol, 2021. **2253**: p. 175-183.
197. Julian, M.C., et al., Efficient affinity maturation of antibody variable domains requires co-selection of compensatory mutations to maintain thermodynamic stability. Sci Rep, 2017. **7**: p. 45259.
198. Koenig, P., et al., Mutational landscape of antibody variable domains reveals a switch modulating the interdomain conformational dynamics and antigen binding. Proc Natl Acad Sci U S A, 2017. **114**(4): p. E486-E495.

199. Shirai, H., A. Kidera, and H. Nakamura, H3-rules: identification of CDR-H3 structures in antibodies. FEBS Lett, 1999. **455**(1-2): p. 188-97.
200. Weitzner, B.D., R.L. Dunbrack, Jr., and J.J. Gray, The origin of CDR H3 structural diversity. Structure, 2015. **23**(2): p. 302-11.
201. Tsumoto, K., et al., Role of salt bridge formation in antigen-antibody interaction. Entropic contribution to the complex between henegg white lysozyme and its monoclonal antibody
HyHEL10. J Biol Chem, 1996. **271**(51): p. 32612-6.
202. Arslan, M., et al., Engineering of conserved residues near antibody heavy chain complementary determining region 3 (HCDR3) improves both affinity and stability. Biochim Biophys Acta Proteins Proteom, 2023. **1871**(4): p. 140915.
203. Fernandez-Escamilla, A.M., et al., Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol, 2004. **22**(10): p. 1302-6.
204. Beerten, J., et al., WALTZ-DB: a benchmark database of amyloidogenic hexapeptides. Bioinformatics, 2015. **31**(10): p. 1698-700.
205. Walsh, I., et al., PASTA 2.0: an improved server for protein aggregation prediction. Nucleic Acids Res, 2014. **42**(Web Server issue): p. W301-7.
206. Conchillo-Sole, O., et al., AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics, 2007. **8**: p. 65.
207. Zambrano, R., et al., AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures. Nucleic Acids Res, 2015. **43**(W1): p. W306-13.
208. Sormanni, P., et al., Rapid and accurate in silico solubility screening of a monoclonal antibody library. Sci Rep, 2017. **7**(1): p. 8200.
209. Sormanni, P., F.A. Aprile, and M. Vendruscolo, The CamSol method of rational design of protein mutants with enhanced solubility. J Mol Biol, 2015. **427**(2): p. 478-90.
210. Krissinel, E. and K. Henrick, Inference of macromolecular assemblies from crystalline state. J Mol Biol, 2007. **372**(3): p. 774-97.
211. Arslan, M., D. Karadag, and S. Kalyoncu, Conformational changes in a Vernier zone region: Implications for antibody dual specificity. Proteins, 2020. **88**(11): p. 1447-1457.
212. Adamczak, R., A. Porollo, and J. Meller, Combining prediction of secondary structure and solvent accessibility in proteins. Proteins, 2005. **59**(3): p. 467-75.
213. Wu, T.T. and E.A. Kabat, An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. J Exp Med, 1970. **132**(2): p. 211-50.
214. Dunbar, J. and C.M. Deane, ANARCI: antigen receptor numbering and receptor classification. Bioinformatics, 2016. **32**(2): p. 298-300.
215. Lucas, X., et al., A thorough anion-pi interaction study in biomolecules: on the importance of cooperativity effects. Chem Sci, 2016. **7**(2): p. 1038-1050.
216. Wong, S.E., B.D. Sellers, and M.P. Jacobson, Effects of somatic mutations on CDR loop flexibility during affinity maturation. Proteins, 2011. **79**(3): p. 821-9.
217. D'Angelo, S., et al., Many Routes to an Antibody Heavy-Chain CDR3: Necessary, Yet Insufficient, for Specific Binding. Front Immunol, 2018. **9**: p. 395.
218. Masuda, K., et al., The role of interface framework residues in determining antibody V(H)/V(L) interaction strength and antigen-binding affinity. FEBS J, 2006. **273**(10): p. 2184-94.
219. Tan, P.H., B.M. Sandmaier, and P.S. Stayton, Contributions of a highly conserved VH/VL hydrogen bonding interaction to scFv folding stability and refolding efficiency. Biophys J, 1998. **75**(3): p. 1473-82.
220. Gvritishvili, A.G., A.V. Gribenko, and G.I. Makhatadze, Cooperativity of complex salt bridges. Protein Sci, 2008. **17**(7): p. 1285-90.

221. Shehata, L., et al., Affinity Maturation Enhances Antibody Specificity but Compromises Conformational Stability. Cell Rep, 2019. **28**(13): p. 3300-3308 e4.
222. Tabasinezhad, M., et al., Trends in therapeutic antibody affinity maturation: From in-vitro towards next-generation sequencing approaches. Immunol Lett, 2019. **212**: p. 106-113.
223. Cherf, G.M. and J.R. Cochran, Applications of Yeast Surface Display for Protein Engineering. Methods Mol Biol, 2015. **1319**: p. 155-75.
224. Chao, G., et al., Isolating and engineering human antibodies using yeast surface display. Nat Protoc, 2006. **1**(2): p. 755-68.
225. Zupancic, J.M., et al., Directed evolution of potent neutralizing nanobodies against SARS-CoV- 2 using CDR-swapping mutagenesis. Cell Chem Biol, 2021. **28**(9): p. 1379-1388 e7.
226. Cregg, J.M., et al., Recombinant protein expression in Pichia pastoris. Mol Biotechnol, 2000. **16**(1): p. 23-52.
227. Jacobs, P.P., et al., Pichia surface display:display of proteins on the surface of glycoengineered Pichia pastoris strains. Biotechnol Lett, 2008. **30**(12): p. 2173-81.
228. Lin, S., et al., A novel fragment of antigen binding (Fab) surface display platform using glycoengineered Pichia pastoris. J Immunol Methods, 2012. **375**(1-2): p. 159-65.
229. Khasa, Y.P., et al., Isolation of Pichia pastoris PIR genes and their utilization for cell surface display and recombinant protein secretion. Yeast, 2011. **28**(3): p. 213-26.
230. Andreu, C. and M.L. Del Olmo, Yeast arming systems: pros and cons of different protein anchors and other elements required for display. Appl Microbiol Biotechnol, 2018. **102**(6): p. 2543-2561.
231. Yang, N., et al., The contribution of Pir protein family to yeast cell surface display. Appl Microbiol Biotechnol, 2014. **98**(7): p. 2897-905.
232. Ko, S., Jung, S.T., , Engineering antibodies for dual specificity and enhanced potency. Biotechnol. Bioprocess Eng., 2015. **20**(2): p. 201-210.
233. Chames, P. and D. Baty, Bispecific antibodies for cancer therapy: the light at the end of the tunnel? MAbs, 2009. **1**(6): p. 539-47.
234. Kaplon, H., et al., Antibodies to watch in 2023. MAbs, 2023. **15**(1): p. 2153410.
235. Sedykh, S.E., et al., Bispecific antibodies: design, therapy, perspectives. Drug Des Devel Ther, 2018. **12**: p. 195-208.
236. Eigenbrot, C. and G. Fuh, Two-in-One antibodies with dual action Fabs. Curr Opin Chem Biol,
2013. **17**(3): p. 400-5.
237. Bostrom, J., et al., Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site. Science, 2009. **323**(5921): p. 1610-4.
238. Jiang, G., et al., Evaluation of semi-homogeneous assay formats for dual-specificity antibodies. J Immunol Methods, 2013. **387**(1-2): p. 51-6.
239. Koenig, P., et al., Deep Sequencing-guided Design of a High Affinity Dual Specificity Antibody to Target Two Angiogenic Factors in Neovascular Age-related Macular Degeneration. J Biol Chem, 2015. **290**(36): p. 21773-86.
240. Koenig, P., et al., Tuning the specificity of a Two-in-One Fab against three angiogenic antigens by fully utilizing the information of deep mutational scanning. MAbs, 2017. **9**(6): p.
959-967.
241. Schaefer, G., et al., A two-in-one antibody against HER3 and EGFR has superior inhibitory activity compared with monospecific antibodies. Cancer Cell, 2011. **20**(4): p. 472-86.
242. Dondelinger, M., et al., Understanding the Significance and Implications of Antibody Numbering and Antigen-Binding Surface/Residue Definition. Front Immunol, 2018. **9**: p. 2278.
243. Wu, H., et al., Humanization of a murine monoclonal antibody by simultaneous optimization of framework and CDR residues. J Mol Biol, 1999. **294**(1): p. 151-62.

244. Haidar, J.N., et al., A universal combinatorial design of antibody framework to graft distinct CDR sequences: a bioinformatics approach. Proteins, 2012. **80**(3): p. 896-912.
245. Alsaab, H.O., et al., PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome. Front Pharmacol, 2017. **8**: p. 561.
246. Socinski, M.A., et al., Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. N Engl J Med, 2018. **378**(24): p. 2288-2301.
247. Lee, W.S., et al., Combination of anti-angiogenic therapy and immune checkpoint blockade normalizes vascular-immune crosstalk to potentiate cancer immunity. Exp Mol Med, 2020. **52**(9): p. 1475-1485.
248. Topalian, S.L., C.G. Drake, and D.M. Pardoll, Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell, 2015. **27**(4): p. 450-61.
249. Carsten Pusch, H.S.a.N.B., Increased cloning efficiency by cycle restriction - ligation (CRL). Technical Tips Online, 1997. **2**: p. 35-37.
250. Boone, M., et al., Massively parallel interrogation of protein fragment secretability using SECRiFY reveals features influencing secretory system transit. Nat Commun, 2021. **12**(1): p. 6414.

Ethical approval

İZMİR BİYOTIP VE GENOM MERKEZİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR ETİK KURULU (İBG-GOEK)
KARARI

Toplantı Tarihi	: 26/02/2021	**Toplantı Günü**	: Cuma
Toplantı Sayısı	: 3	**Toplantı Saati**	: 10:30

Sayın Doç. Dr. Sinan GÜVEN,

2021-008 Protokol No'lu; sorumlusu olduğunuz "Antikor mühendisliği yaklaşımları ile iyileştirilmiş özelliklere sahip anti-kanser antikor parçacıklarının geliştirilmesi" başlıklı araştırmanın uygulanmasında etik açıdan sakınca olmadığına oy birliği ile karar verilmiştir.

Bilgilerinizi ve gereğini rica ederiz.

Prof. Dr. H. Alper BAĞRIYANIK
Başkan

Pandemi süresince online ortamda gerçekleştirilen toplantımızda alınan kararlar tek imzalı olarak düzenlenmektedir. Pandemi sona erdikten sonra ıslak imzalı karar belgesi teslim edilecektir.

Cirruculum vitae

MERVE ARSLAN

PERSONAL DATA

Address : Aydoğdu mah. Yavuz Sultan Selim Cad. No: 22, 59200 Süleymanpaşa, Tekirdağ, Türkiye
Phone : +90 5388276672
E-mail : amerve26@gmail.com

EDUCATION

Feb 2018 – Jun 2023 : Integrated Ph.D. Student in Molecular Biology and Genetics, Izmir International Biomedicine and Genome Institute, Dokuz Eylul University, Izmir, Turkey

Jun 2022– Dec 2022 : Visiting Ph.D. Student, Callewaert Lab, VIB-UGent, Center for Medical Biotechnology, Ghent University-Flanders Institute for Biotechnology, Ghent, Belgium

Sept 2012 – Jul 2017 : Bachelor's Degree in Bioengineering, Faculty of Engineering, Ege University, Izmir, Turkey

Feb 2016– Jul 2016 : Socrates-Erasmus Exchange Programme in Department of Biotechnology, University of Science and Technology UTP, Bydgoszcz, Poland

RESEARCH EXPERIENCES

Doctoral Research Assistantship

Jun 2018 – Jul 2022 : Antibody Engineering Laboratory, Izmir Biomedicine and Genome Center, IBG, Izmir, Turkey, "Development of New Generation Recombinant Antibody Fragments Binding to VEGF for Cancer Treatment", *119Z161-TUBITAK*, Advisor : Sibel Kalyoncu

May 2020 – Feb 2021 : IBG-Pharma Microbial Platform, Izmir Biomedicine and Genome Center, IBG, Izmir, Turkey, "Development of SARS-CoV-2 Vaccine Based on Recombinant Protein Applications",*TUBITAK* Advisor : Mehmet Inan

SCHOLARSHIPS

Jun 2020 – Jun 2023 : The Scientific And Technological Research Council Of Turkey TUBITAK-BIDEB – 2211A Government Achievement Grant

Feb 2018 – Mar 2023 : Council of Higher Education (YOK), Integrated Ph.D scholarship within 100/2000 Doctorate Scholarship

Sept 2022 – Dec 2022 : The Scientific And Technological Research Council Of Turkey TUBITAK-BIDEB – 2214A International Research Fellowship Programme for PhD Students

Jun 2022 – Sept 2022 : European Molecular Biology Organization (EMBO) Sientific Exchange Grants (9431)

Oct 2019 – Jun 2020 : The Scientific And Technological Research Council Of Turkey TUBITAK-ARDEB – Research Fellowship (119Z161) **Feb 2016 – Jul 2016 :** Socrates – Erasmus Exchange Student Scholarship

Publication List

Merve Arslan, Tuğçe Uluçay, Seyit Kale, Sibel Kalyoncu, Engineering of conserved residues near antibody heavy chain complementary determining region 3 (HCDR3) improves both affinity and stability, Biochimica et Biophysica ACTA (BBA) – Proteins and Proteomics, 2023 April, doi:10.1016/j.bbapap.2023.140915.

Sibel Kalyoncu, Semiramis Yilmaz, Ayca Zeybek Kuyucu, Dogu Sayili, Olcay Mert, Hakan Soyturk, Seyda Gullu, Huseyin Akinturk, Erhan Citak, **Merve Arslan,** Melda Guray Taskinarda, Ibrahim Oguzhan Tarman, Gizem Yilmazer Altun, Ceren Ozer, Ridvan Orkut, Aysegul Demirtas, Idil Tilmensagir, Umur Keles, Ceren Ulker, Gizem Aralan, Yavuz Mercan, Muge Ozkan, Hasan Onur Caglar, Gizem Arik, Mehmet Can Ucar, Muzaffer Yildirim, Tugce Canavar Yildirim, Dilara Karadag, Erhan Bal, Aybike Erdogan, Serif Senturk, Serdar Uzar, Hakan Enul, Cumhur Adiay, Fahriye Sarac, Arzu Tas Ekiz 1, Irem Abaci, Ozge Aksoy, Hivda Ulbegi Polat, Saban Tekin, Stefan Dimitrov, Aykut Ozkul, Gerhard Wingender, Ihsan Gursel, Mehmet Ozturk, Mehmet Inan, Process development for an effective COVID-19 vaccine candidate harboring recombinant SARS-Cov-2 delta plus receptor binding domain produced by Pichia pastoris, Scientific Reports, 2023 March, doi:10.1038/s41598-023-32021-9.

Merve Arslan, Murat Karadag, Ebru Onal, Emine Gelinci, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Effect of non-repetitive linker on *in vitro* and *in vivo* properties of an anti-VEGF scFv, Scientific Reports, 2022 March, doi:10.1038/s41598-022-09324-4.

Deniz Doğan**, Merve Arslan,** Tuğçe Uluçay, Sibel Kalyoncu, Stefan Dimitrov, Seyit Kale, CENP-A Nucleosome is a Sensitive Allosteric Scaffold for DNA and Chromatin Factors, Journal of Molecular Biology, 2021 March, doi:10.1016/j.jmb.2020.166789

Merve Arslan, Dilara Karadag, Sibel Kalyoncu, Conformational changes in a Vernier zone region: implications for antibody dual specificity, Proteins: Structure, Function, and Bioinformatics, 2020 June, doi.org:10.1002/prot.25964.

Merve Arslan, Dilara Karadag, Sibel Kalyoncu, Protein engineering approaches for antibody fragments: directed evolution and rational design approaches. Turkish Journal of Biology, 2019 February; 43 (1): 1-12. doi:10.3906/biy-1809-28.

Merve Arslan, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Antibody engineering approaches on Vernier zone residues to develop biobetter antibody therapeutics for cancer treatment, Proceeding of the 2nd International Ion Channel and Cancer Congress (IonCC2019), Turkish Journal of Biochemistry, 44(s2), 24-64. doi:10.1515/tjb-2019-44s206

Book Chapters

Merve Arslan, Dilara Karadag, Sibel Kalyoncu, Antikorlar İçin Protein Mühendisliği Yaklaşımları, Protein: Yapisi, Mühendisliği, Etkileşimleri, Dinamiği Ve İlaç Tasarimindaki Yeri, Ankara Nobel Tıp Yayınevi, 2021.

Merve Arslan*, Nazlı Eda Kaleli *, Murat Karadağ*, Murat Caner Yarımçam*, Sibel Kalyoncu, Hanin Alotaibi, Kodon Optimizasyonu Ve Rekombinant Protein Üretimindeki Yeri, Protein: Yapisi, Mühendisliği, Etkileşimleri, Dinamiği Ve İlaç Tasarimindaki Yeri, Ankara Nobel Tıp Yayınevi, 2021.

Murat Karadag, **Merve Arslan**, Nazlı Eda Kaleli, Sibel Kalyoncu, Physicochemical determinants of antibody-protein interactions. APCSB: Protein Interactions as Targets in Drug Discovery, 2019 November; 121:85-114, Elsevier. doi:10.1016/bs.apcsb.2019.08.011

Selected Abstracts

Merve Arslan, Murat Karadag, Ebru Onal, Emine Gelinci, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Effects of single chain variable fragment (scFv) antibody linkers on *in vitro* and *in vivo* properties, BioTurkey International Biotechnology Virtual Congress, 9-11 Sept 2021 – Poster presentation.

Merve Arslan, Tugce Ulucay, Seyit Kale, Sibel Kalyoncu, Co-evolution of antibody affinity and stability: HCDR3 reveal its role in affinity/stability trade-offs, European Molecular Biology Organisation (EMBO) Virtual Practical Course, Integrative modelling of biomolecular interactions, 30 May – 5 June 2021 - Poster presentation.

Merve Arslan, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Mutation on heavy chain V112 highly affects the thermal stability of the antibody fragments, V. Turkey *in vitro* Diagnostic Symposium, 19-21 Feb 2020, Izmir, Turkey – Oral presentation.

Merve Arslan, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Development of anti-angiogenesis antibody fragments for use in new generation cancer treatments: The importance of Vernier regions in antibody engineering, "Onkolojide İz Bırakanlar Zirvesi", 14-19 Nov 2019, Antalya, Turkey – Oral presentation.

Merve Arslan, Dilara Karadag, Sibel Kalyoncu, Spatial positioning of LCDR1 and Vernier zone residues is important for dual specific antibodies, The International Symposium on Health Informatics and Bioinformatics, (HIBIT2019), 17-18 Oct 2019, Izmir, Turkey – Oral presentation, ***Best Selected Talk Award***

Merve Arslan, Gulcin Cakan-Akdogan, Sibel Kalyoncu, Antibody engineering approaches on Vernier zone residues to develop biobetter antibody therapeutics for cancer treatment, 2nd International Ion Channel and Cancer Congress (IonCC2019), 22-24 Sept 2019, Izmir, Turkey - Oral presentation.

TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Review Article

Turk J Biol
(2019) 43: 1-12
© TÜBİTAK
doi:10.3906/biy-1809-28

Protein engineering approaches for antibody fragments: directed evolution and rational design approaches

Merve ARSLAN[1,2,*], **Dilara KARADAĞ**[1,*], **Sibel KALYONCU**[1,**]
[1]İzmir Biomedicine and Genome Center, İzmir, Turkey
[2]İzmir Biomedicine and Genome Institute, Dokuz Eylül University, İzmir, Turkey

Received: 07.09.2018 • **Accepted/Published Online:** 09.01.2019 • **Final Version:** 07.02.2019

Abstract: The number of therapeutic antibodies in preclinical, clinical, or approved phases has been increasing exponentially, mostly due to their known successes. Development of antibody engineering methods has substantially hastened the development of therapeutic antibodies. A variety of protein engineering techniques can be applied to antibodies to improve their affinity and/or biophysical properties such as solubility and stability. Antibody fragments (where all or some parts of constant regions are eliminated while the essential antigen binding region is preserved) are more suitable for protein engineering techniques because there are many in vitro screening technologies available for antibody fragments but not full-length antibodies. Improvement of biophysical characteristics is important in the early development phase because most antibodies fail at the later stage of development and this leads to loss of resources and time. Here, we review directed evolution and rational design methods to improve antibody properties. Recent developments in rational design approaches and antibody display technologies, and especially phage display, which was recently awarded the 2018 Nobel Prize, are discussed to be used in antibody research and development.

Key words: Antibody, antibody fragment, directed evolution, rational design, protein engineering, phage display, yeast surface display, affinity, biophysical properties

1. Introduction

Hundreds of therapeutic antibodies and their derivatives are being manufactured and tested in clinical trials. Currently, there are more than 65 monoclonal antibodies approved on the market for the treatment of various diseases, mostly cancer. The rate of antibody therapeutics receiving their first approvals has been increasing over the last decade. Last year, 10 antibodies were approved in either the European Union or the United States and this number is expected to increase in the upcoming years (Kaplon and Reichert, 2018).

The first technology that was used to produce therapeutic antibodies was mouse hybridoma technology (Frenzel et al., 2017). With this technology, therapeutic monoclonal antibodies (mAbs) are obtained via the fusion of murine B cells and myeloma cells. However, there are some limitations in the use of these mAbs in humans, especially the immune response against murine mAbs (human antimouse antibody response) (Qin and Li, 2014). To overcome this problem, several approaches were developed by utilizing recombinant DNA technology, such as chimerization (replacement of the constant regions of the murine antibodies with homologous human sequences), which generally reduces the affinity and deteriorates biophysical properties of mAbs. Therefore, it is essential to apply affinity maturation and protein engineering approaches a%er this process. More importantly, there are known reproducibility problems related to the hybridoma technique where sequence information is lost and features of mAbs cannot be improved with many available in vitro systems (Bradbury and Pluckthun, 2015).

Approximately 90% of approved antibody drugs are full-length (IgG) and the rest are antibody fragments (mostly Fab formats), where all or some parts of constant regions are eliminated while the essential antigen binding region is preserved. It is very well known that antibody fragments usually show similar binding properties as their full-length versions with even better biophysical properties (Nelson, 2010). Compared to full-length antibodies, antibody fragments have many advantages for therapeutic

* These authors contributed equally to this work
** Correspondence: sibel.kalyoncu@ibg.edu.tr

use: (i) lower immunogenicity due to lack of constant regions, (ii) higher tumor penetration, (iii) cheaper and larger scale production with bacteria, and (iv) availability of various in vitro display technologies to improve several characteristics of antibodies. Today, the number of antibody fragments in clinical trials and on the market is increasing faster than before due to their advantages. Because most of the directed evolution approaches are only available for antibody fragments, improvement of full-length antibodies is usually conducted in their antibody fragment format, and then those improved fragments are converted back to full-length antibody format (Xiao et al., 2017).

Protein engineering techniques such as directed evolution and rational design approaches to discover and/or improve antibodies are becoming more popular both in the biopharmaceutical industry and research environments. Applying these techniques in the early discovery phase is important because it is high-throughput and there is full control of protein sequence during the development phase of biotherapeutics.

2. Antibody display technologies as directed evolution approaches

For the past 40 years, hybridoma technology has been used extensively to produce traditional monoclonal antibodies for research and diagnostics. Recently, a number of advanced methods called display technologies have emerged as fast and high-throughput alternatives. Phage display technology is the first radical in vitro approach that allowed to produce human antibodies without any need for immunization. In this technique, antibody fragments are fused to a capsid protein of the phage and thus expressed on the surface of the virus (García Merino, 2011; Chiu and Gilliland, 2016). Although phage display is the most common antibody display technique, today several recombinant display technologies are available and basically classified in two categories: in vitro display technologies (phage display, ribosome-mRNA display) and in vivo display technologies (bacterial, yeast, and mammalian cell-surface display) (Sergeeva et al., 2006; Harel Inbar and Benhar, 2012; Brodel et al., 2018).

2.1. In vitro display technologies

2.1.1. Phage Display

The phage display technique was first discovered in 1985 by George P Smith, who was one of three recipients of the 2018 Nobel Prize in chemistry for this discovery (Smith, 1985). This was an important step to develop new approaches for generation of mAbs. In this technique, a protein gene is fused to a gene encoding a capsid protein of the virus and the fused gene is inserted into a single-stranded DNA of the phage (Karimi et al., 2016; Ledsgaard et al., 2018). Basically, two types of capsid proteins are preferred; the first one is pIII that allows to fuse larger proteins and the other one is pVIII. The most commonly used phages for phage display are the filamentous ones (M13, Fd, and f1), which are in the Ff family and have the ability to infect only the strains of *Escherichia coli* containing F conjugative plasmids (Li and Caberoy, 2010; Loset and Sandlie, 2012; Karimi et al., 2016; Gustafson et al., 2018; Kiguchi et al., 2018; Ledsgaard et al., 2018; Teixeira and Gonzalez-Pajuelo, 2018).

Two different application systems have been developed for phage display. In the first system, a protein sequence is used as an insert and is fused to a capsid gene of the virus. With this system, the desired protein is expressed within the genome of the virus. In the other more preferred system, a different plasmid called a phagemid is used and the expression of the desired protein is separated from the phage replication. Phagemids also include replication origins of *E. coli* and a phage, a specific selection marker, and specific tags that help detection and purification of the desired protein (Li and Caberoy, 2010; Loset and Sandlie, 2012; Teixeira and Gonzalez-Pajuelo, 2018).

The phage display technique was first applied for the variable fragments of antibodies and many different antibody fragment formats have been displayed by this technique. The antibody fragments that are displayed by this technique are usually scFv (a single-chain variable fragment) or Fab (antigen-binding fragment), and nowadays the most popular ones are V_H (nanobody, heavy variable domain of the antibody) (Teixeira and Gonzalez-Pajuelo, 2018). It is easy to convert these fragments to full-length antibodies by recombinant DNA technology, if needed.

The phage display technique is carried out by a process of in vitro repeated cycles typically named biopanning or phage display selections (Figure 1). This process includes the following steps: (1) incubation: binding of the antibody library repertoire to the antigen; (2) washing: elimination of the nonspecific binders; and (3) elution and amplification: obtaining antibodies binding to the antigen specifically for further cycles or for screening. Although in the first cycle of biopanning the whole antibody repertoire is exposed to the antigen, depending on the fragment type of the antibody and the phage display, 2–4 cycles of selections are generally performed to enrich the specific binders. Evaluation of the success of each cycle of the process and enrichment is possible by comparing the phage titers after elution steps against a blank that does not include antigen, or alternatively it can be tested by ELISA (enzyme-linked immunosorbent assay) (Hairul Bahara et al., 2013; Chan et al., 2014; Ledsgaard et al., 2018; Teixeira and Gonzalez-Pajuelo, 2018).

The most powerful advantages of phage display are its small size and high diversity (antibody libraries up to

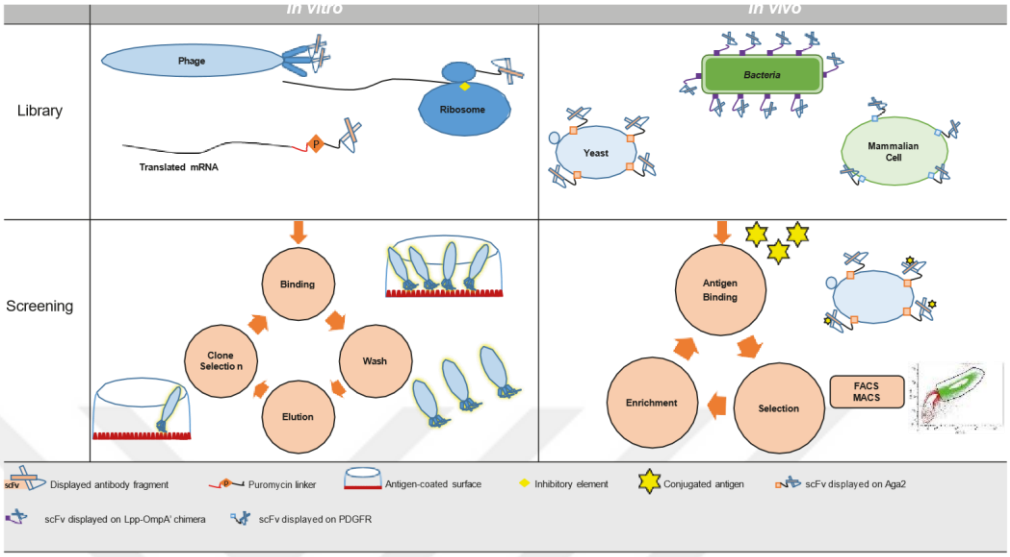


Figure 1. Antibody display technologies. General schematic for in vitro and in vivo display techniques.

10^{11} clones), which allow to obtain antibody fragments with the desired affinity and biophysical properties. Also, this technique is preferred in both research areas and the biopharmaceutical industry due to its library diversity, ease of use, and low cost (Liu et al., 2017; Teixeira and Gonzalez-Pajuelo, 2018). For example, belimumab (market name Benlysta) was discovered and improved by phage display and it is used to treat adults with active systemic lupus erythematosus (Stohl and Hilbert, 2012). As a better known example, adalimumab (market name Humira) was discovered by phage display and it is widely used for rheumatoid arthritis treatment (Bain and Brazil, 2003). The number of antibodies discovered and/or optimized by phage display has been exponentially increasing over the last decade due to the many advantages listed above (Nixon et al., 2014).

2.1.2. Ribosome and mRNA display

Ribosome and mRNA display techniques are cell-free and this feature separates them from other display platforms. They have high molecular diversity (antibody libraries ranging from 10^{12} to 10^{14} clones) and enable isolation of antibodies that show affinities at pM level. Both techniques include the same basic features, such as in vitro transcription and translation steps (Harel Inbar and Benhar, 2012).

Ribosome-display technology was first reported in a patent application in 1991. While the mRNA encoding antibody library is translated in vitro, the translated peptide and corresponding mRNA remain attached to the ribosome (Figure 1). By this means, the peptide–ribosome–mRNA (PRM) complex can be selected along with the sequence information of the desired antibody by affinity purification techniques. The most powerful aspect of this technique is its large size of library, which is not limited by the cell transformation efficiency. On the other hand, the ribosome amount and the existence of unrelated mRNA molecules are the main limitations of this technique (Hanes and Pluckthun, 1997). Groves et al. compared phage display and ribosome display to generate scFvs to a specific antigen (Groves et al., 2014). They found that scFvs affinity-matured by ribosome display had more structural diversity in the HCDR3 and V_H-V_L interface regions.

In the mRNA-display technique, first the antibody DNA library is transcribed to mRNA. Then mRNA is ligated to a linker, which is a DNA sequence linked to puromycin. Thereafter, the mRNA-linker-puromycin complex is translated. Puromycin first binds to the A-site of the ribosome, then attacks the P-site and the nascent peptide is transferred to puromycin, resulting in the mRNA-linker-puromycin-antibody fragment complex. The complex is then reverse-transcribed and the selection process is performed. After the selection step, ss-DNA is obtained by hydrolyzing the complementary mRNA via high pH, and the desired DNA sequence is amplified by PCR (Figure 1) (Sergeeva et al., 2006; Jijakli et al., 2016; Liu et al., 2017). One of the limitations of this technique is low efficiency of mRNA-protein conjugates. Nagumo et al. overcame this problem by unexpected substitution mutations around the start codon of antibodies (Nagumo et al., 2016). These mutations destabilized the mRNA secondary structure and this somehow led to a better formation of conjugates and higher protein expression.

2.2. In vivo display technologies

2.2.3. Mammalian surface display

10^9 copies of scFv, which is a limitation as compared with the other display platforms such as phage display (Chao et al., 2006; Harel Inbar and Benhar, 2012). Also, the best known disadvantage of yeast surface display is slower growth rate and lower transformation efficiency compared to both phage and bacteria surface display techniques (Mei et al., 2017).

2.2.1.

The bacterial surface display technique was developed as a potential alternative to phage display (Sergeeva et al., 2006). The use of bacteria as a display system was first reported by George Georgiou's group in 1993 (Georgiou et al., 1993). They first used the Lpp-OmpA' chimera to display two specific scFvs on the outer membrane of the gram-negative bacterium *E. coli*. Several years later, a new approach was developed by the same group called APEx (anchored periplasmic expression) (Jeong et al., 2007). With this second system, scFvs were displayed in the periplasmic space anchored to the inner membrane of *E. coli*. For isolation of antigen-specific clones, flow cytometry was used for both applications. Due to the technological shortcoming of the FACS (fluorescence activated cell sorting) of that time, library size was limited and thereby this technique was basically used for the evolution of the preexisting antibodies (Harel Inbar and Benhar, 2012). This technique is more commonly used to display functional enzymes, antigens, and especially polypeptide libraries (up to 10^{11} library size) (Sergeeva et al., 2006; Liu et al., 2017). For example, to identify peptide ligands specific for VEGF, bacteria-displayed peptide libraries were constructed and screened (Liu et al., 2017).

2.2.2. Yeast surface display

Yeast surface display was first demonstrated by Dane Wittrup's group using *Saccharomyces cerevisiae* to display antibody repertoires (Harel Inbar and Benhar, 2012). Yeast surface display is a powerful technique that allows to obtain antibodies with desired affinity, specificity, and stability. In this technique, scFvs that consist of V_H and V_L regions and a polypeptide linker binding them together are displayed. On yeast, scFvs are fused to the adhesion subunit of the yeast agglutinin protein Aga2p, which is bound to Aga1p via a disulfide bond and this complex attaches the scFv to the yeast cell wall and finally the desired antibody fragment is identified by FACS (Figure 1) (Feldhaus and Siegel, 2004; Chao et al., 2006; Liu et al., 2017; Mei et al., 2017). This technique is commonly used for antibody display and it has several advantages: (i) use of FACS to monitor equilibrium activity statistics of the sample; (ii) o%ering easy secretion and purification; and (iii) using yeast cells, which can perform posttranslational modification. On the other hand, it allows to display up to

Bacterial surface display

4

Mammalian surface display was developed by Ira Pastan's group in 2006 (Ho et al., 2006). They used this technique to display an scFv library fused to the N-terminal transmembrane domain of human platelet-derived growth factor receptor (PDGFR) on the surface of HEK-293T cells and were able to isolate high-affinity anti-CD22 antibodies. Mammalian cell display has powerful aspects for the isolation of scFv and whole IgG with high affinity and other specific biological functions. For instance, they can express mouse or human antibodies containing the posttranslational modifications required for some key antibody functions, and the technique can also be used to express recombinant antibody fragments that cannot be expressed in *E. coli* (Ho and Pastan, 2009). However, there are only a few reports of the technique, basically due to the limitation of repertoire size (ranging between 10^3 and 10^6). Similar to other techniques, it is required to transfer genes encoding the desired proteins to proper host cells by convenient vectors and to make sure that the desired protein undergoes correct transcription and translation processes. However, due to slower proliferation rates of mammalian cells in contrast to microbial ones, it is challenging to choose cells suitable for construction of a convenient and rapid mammalian cell surface display system. HEK-293, COS, and CHO cells are the most widely used cell lines in mammalian surface display approaches. HEK-293 has particularly been preferred more than others because of its ease of transport, high yield, and native human glycosylation (Qin and Li, 2014).

3. Rational design approaches

Aggregation, solubility, and stability are important factors that a%ect the developability of an antibody. These challenges can occur during the production process due to the protein's large complex profile and can cause reduced antigen binding affinity, immunogenic responses, and waste of resources. Aggregation/solubility and stability properties of an antibody depend on both its sequence and structure (Figure 2). It is advantageous to control these properties with rational design before in vitro and in vivo studies. Rational design methods aim to demonstrate problematic regions of protein sequences or structures. Thus, combining rational design methods and in vitro/in vivo studies enhances the chance of antibodies with better solubility and stability in the early production phase.

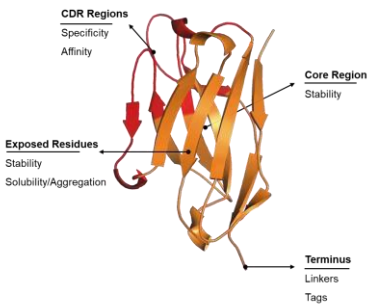


Figure 2. Protein engineering approaches based on antibody regions. While solubility/aggregation can be improved by engineering exposed residues, stability can be increased by both exposed and core residues. Complementarity-determining regions (CDRs) mainly affect specificity and affinity. Tags/linkers can be added for better functionality.

Intrinsic and extrinsic properties play important roles in rational design predictions (Dubay et al., 2004; Pawar et al., 2005). Physicochemical properties of amino acids affect the antibody profile as intrinsic factors. For example, the aggregation rate of the polypeptide can be increased when the number of hydrophobic residues increases. Also, extrinsic factors such as pH, ionic strength, and temperature should be considered during sequence-based prediction (Dubay et al., 2004). Both intrinsic and extrinsic factors might change the properties and only aggregation/solubility can be predicted based on the sequence because the aggregation rate usually depends on amino acid properties. On the other hand, stability and affinity depend on both amino acid and structure properties.

Recent developments in rational design tools that predict problematic regions of proteins help researchers to improve antibody developability. Here, we introduce rational design web tools that can be used to improve aggregation/solubility, stability, and affinity properties based on mutagenesis.

3.1. Aggregation/solubility

Protein aggregation is a common problem in therapeutic antibodies and it can occur during production or storage. On the molecular level, aggregation occurs due to specific regions of a protein sequence named aggregation prone regions (APRs) that determine its aggregation rate. These APRs indicate specific charge, hydrophobicity, and secondary structural properties and lead to aggregation (Fink, 1998; Tartaglia and Vendruscolo, 2008; Agrawal et al., 2011; Elgundi et al., 2017). Prediction of potential APRs is the key function of aggregation/solubility prediction tools.

Early studies showed that protein aggregation and stability kinetics are computable and protein sequences can be designed based on desired properties (Kamtekar et al., 1993; West et al., 1999; Worn and Pluckthun, 1999; Worn and Pluckthun, 2001). Several prediction tools have been developed to determine the aggregation propensity of a protein (Table). While most of them analyze the amyloid formation, some analyze only aggregation propensity/APRs. However, most of the tools can be used to analyze antibody fragments due to their small size. The most commonly used tools are Tango (Fernandez-Escamilla et al., 2004), Waltz (Beerten et al., 2015), AggreScan (Conchillo-Sole et al., 2007), Pasta 2.0 (Walsh et al., 2014), and Camsol Instrinsic (Sormanni et al., 2015), which determine the aggregation propensity of an antibody based on its sequence.

Tango (http://tango.switchlab.org/) is the earliest aggregation prediction tool and predicts the β-sheet aggregation of a given protein sequence. It evaluates probability scores for each amino acid's beta turn, beta

Table. Protein sequence and structure-based web tools for rational design approaches.

Sequence-based Prediction Web Tools			
	Tool name	Definition	References
Aggregation/Solubility	Pasta 2.0	Predicts aggregation-prone, disordered regions	(Walsh et al., 2014)
	Tango	Evaluates the aggregation scores for each residue based on physicochemical principles	(Fernandez-Escamilla et al., 2004)
	Waltz	Computes the position-specific scores to determine aggregation-prone regions	(Beerten et al., 2015)
	Camsol Intrinsic	Gives output scores for each residue based on solubility profile of sequence	(Sormanni et al., 2015)
	AggreScan	Predicts aggregation-prone regions and estimates the effect of mutation on aggregation profile	(Conchillo-Sole et al., 2007)
	Fish Amyloid	Identifies amyloidogenic regions in protein sequences	(Gasior and Kotulska, 2014)
	Soda	Focuses on effect of the mutations on intrinsic solubility profile of protein sequences	(Paladin al., 2017)
	Pon-Sol	Determines the effect of amino acid variation on solubility profile	(Yang al., 2016)
	Protein-Sol	Gives graphical outputs of highlighted lysine arginine contents and solubility profile	(Hebditch et al., 2017)
Structure Based Prediction Web-Tools			
Aggregation/solubility	Camsol Structurally Corrected	Gives structurally corrected solubility profile to visualize poorly soluble regions on the surface, determines the proper residues for mutation	(Sormanni et al., 2015)
	AggreScan3D	Identifies the poorly soluble residues based on both position of amino acid and amino acid structure	(Zambrano et al., 2015)
Stability	ProMaya	Calculates protein stability based on differences between protein's wild type and mutated type free energies	(Wainreb et al., 2011)
	CDM	Evaluates the stability differences between the wild type and mutated type protein structure	(Pandurangan et al., 2017)
	Mutant	Determines the structure and sequence-based stability changes depending on single point mutation of protein	(Capriotti et al., 2005)
	Cupsat	Uses amino acid–atom potential and torsion angle distribution information to identify changes in protein stability-based on mutations	(Parthiban et al., 2006)
Affinity	mCSM AB	Predicts antigen antibody affinity changes upon mutations	(Pires and Ascher, 2016)

sheet, alpha helix, and beta and alpha aggregation considering given extrinsic conditions (pH, temperature, ionic strength, concentration). The algorithm assumes that specific regions of protein have high aggregation propensity if they involve at least five consecutive residues with a probability to populate the β-aggregate state higher than 5% per residue. It was shown that Tango has a success rate of 87% , correctly predicting 155 out of 179 peptides, with 21 false positives and 3 false negatives (Fernandez-Escamilla et al., 2004).

Waltz (http://waltz.switchlab.org/) and Pasta 2.0 (http://protein.bio.unipd.it/pasta2/) give highly aggregation-prone/amyloid-forming regions as output. While Waltz uses a position-specific scoring matrix (PSSM) with physicochemical information to identify amyloid forming regions (Beerten et al., 2015), Pasta 2.0 identifies amyloid forming regions by calculating the pairing energies for each pair of residues facing one another on parallel or antiparallel neighboring strands within a β-sheet (Walsh et al., 2014).

AggreScan and Camsol are listed in two subsections of the Table because they can analyze protein aggregation propensity based on both sequence and structure information. Aggrescan (http://bioinf.uab.es/aggrescan/) calculates aggregation propensity scores for each residue in the sequence by averaging the aggregation propensity score per residue over a given length (Conchillo-Sole et al., 2007). Aggrescan3D (A3D) (http://biocomp.chem.uw.edu.pl/A3D/) is an improved version of Aggrescan that overcomes the limitations of sequence-based analyses. A3D identifies aggregation prone residues, which are related to folded states. Also, designed/desired mutation effects on aggregation propensity of any protein can be determined by using A3D (Zambrano et al., 2015).

The Camsol method can be used in two different modes, 'Camsol Intrinsic' and 'Camsol Structurally Corrected' (http://www-vendruscolo.ch.cam.ac.uk/camsolmethod.html), to evaluate aggregation scores of any protein. Camsol Intrinsic calculates the solubility profile scores per amino acid by using the given protein sequence and identifies the regions that are poorly soluble when the score is smaller than –1. It evaluates the aggregation propensity per residue using the sequence, charge, hydrophobicity, and secondary structure propensity as intrinsic factors. Camsol Structurally Corrected analyzes the protein structure like Camsol Intrinsic but it shows the poorly soluble regions on the surface that can be used to identify suitable mutations to increase the solubility of the protein. These poorly soluble regions can also be visualized by using output structure (Sormanni et al., 2015, 2017).

These methods can be used separately or combined to predict aggregation/solubility profiles and the combination of different methods can provide higher accuracy for mutagenesis studies. Van Der Kant et al. used only Tango for prediction of APRs as a part of a study analyzing the relationship between intrinsic aggregation propensity and the local thermodynamic stability of over 2000 antibody structures from the abYsis database (Van Der Kant et al., 2017). Wang et al. combined Tango with structure-based methods to predict APRs in antibody sequences based on 29 published Fab-antigen complexes (Wang et al., 2010). They tested two different thresholds and they found that Tango was more than 92% correct in their experimental validation studies. In another study, estimations of Tango, Aggrescan, and Pasta 2.0 were used to identify APRs that were mostly confirmed by experimental results (Yageta et al., 2015).

Lately several sequence-based aggregation propensity prediction tools have also been developed. Gasior and Kotulska proposed a classification method called Fish Amyloid (http://comprec-lin.iiar.pwr.edu.pl/) that is able to recognize amyloidogenic fragments based on well-defined patterns of residue distribution and cooccurrence of position-specific amino acids in protein sequences (Gasior and Kotulska, 2014). Fish Amyloid was trained on different lengths of sequences and offered good potential for prediction. PonSol (http://structure.bmc.lu.se/PON-Sol) determines the effect of amino acid variations solubility profiles. The tool uses 443 amino acid substitutions from 71 proteins and these amino acid substitutions are classified as increasing, decreasing, and not affecting solubility (Yang et al., 2016). Protein-Sol (https://protein-sol.manchester.ac.uk/) is another recent sequence-based prediction tool that uses datasets of *Escherichia coli* protein solubility for comparison and calculates 35 sequence-based properties. The tool gives graphical output of predicted solubility, fold propensity, and net segment charge. Predicted solubility scales from 0 to 1 and more than 0.45 solubility scores are accepted as soluble. Also, lysine and arginine contents are highlighted for modifying protein solubility (Hebditch et al., 2017). Soda (http://protein.bio.unipd.it/soda/) predicts the protein solubility changes based on calculations of several physicochemical properties for given mutations. The method compares the mutant type and wild type profile properties and estimates the changes. Soda provides convenience for different types of variations such as point mutation, deletion, or insertion (Paladin et al., 2017).

As a case study, an scFv sequence used in our lab was analyzed with some of the sequence-based tools introduced above (Figure 3). The full scFv sequence was given as input. As output, every residue had an aggregation/solubility score based on the tool's calculation and they were highlighted as aggregation-prone according to the tool's corresponding thresholds. We determined multiple regions of the scFv as aggregation-prone (at least 6 of 8 tools gave predicted aggregation-prone residues).

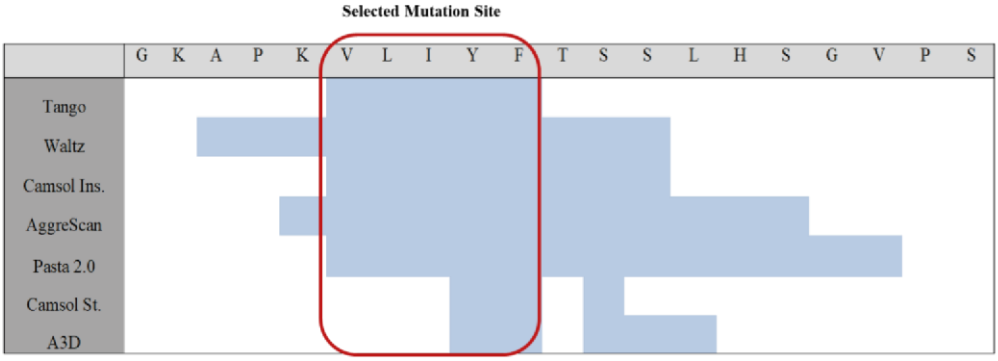


Figure 3. A case study for web tools. Several tools introduced in this review were used to determine aggregation-prone regions of an scFv sequence used in our lab. Blue highlighted regions are outputs of tools as aggregation prone regions. Each web tool has a different threshold, which was not shown in this figure. Mutation site is selected according to common predicted regions of different tools (at least 6 of 8 tools gave same residues as aggregation-prone).

One of those regions is shown as an example in Figure 3. Our future mutations will be focused on those regions to improve the biophysical characteristics of our protein.

3.2. Stability

Protein stability can be predicted by calculating the change in the Gibbs free energy due to substitution of an amino acid and more negative values of free energy present better stability (Thiltgen and Goldstein, 2012). Different approaches can be used for prediction of protein stability, such as physical, statistical, empirical, and/or machine learning methods. While the first three approaches are limited and are more time- and cost-intensive, machine learning methods can quickly perform predictions based on input mutation, protein sequence, and structural information at the same time (Capriotti et al., 2004; Cheng et al., 2006).

Several web-based tools were developed to predict protein stability. ProMaya (http://bental.tau.ac.il/ProMaya/) calculates the stability free energy change upon mutations by combining a collaborative filtering-based algorithm (CF) and random forest regression. The tool uses different available datasets of mutations in the same and different positions. ProMaya suggests that using known free energy values of mutations at a specific position corrects the prediction of free energy differences for other mutations (Wainreb et al., 2011).

SDM (http://marid.bioc.cam.ac.uk/sdm2) evaluates the stability change between the wild type and mutant protein by using a conformationally constrained environment-specific substitution table (ESST). The method analyzes the amino acid alteration with specific structural parameters based on residue packing density and the ESST. The webserver gives predicted stability difference scores interpreted as reduced, induced, or unaffected stability (Pandurangan et al., 2017).

I-Mutant (http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi) predicts protein stability changes based on a support vector machine and allows users to use protein structure or sequences for prediction. It was shown that I-Mutant has an accuracy of 77%–80% for the dataset derived from ProTherm (Bava et al., 2004; Capriotti et al., 2005).

Cupsat (http://cupsat.tu-bs.de/) uses atom potential and torsion angle distribution information of amino acids to identify protein stability free energy change upon mutations. The tool analyzes the protein structure and gives information about mutation site, solvent accessibility, and torsion angle and whether the mutated amino acid has suitable torsion angles or not. It was shown that Cupsat achieved 80% prediction success for both thermal and chemical stability (Parthiban et al., 2006).

3.3. Affinity/specificity

If affinity improvement is desired, in vitro/vivo methods explained Section 2 of this review can be used. There are many available affinity maturation strategies based on directed evolution methods. Generally, mutations in complementarity-determining regions (CDRs) for improving antigen–antibody affinity cannot be predicted by using rational design approaches because it is hard to estimate the dynamic antigen–antibody complex structure. However, there is a newly developed tool called mCSM-AB (http://biosig.unimelb.edu.au/mcsm_ab/) that uses free energy change upon mutation and estimates the affinity change. In the tool, a negative sign means that the selected mutation reduces affinity and a positive sign means that the selected mutation increases affinity. It is important to

know that this tool allows users to select more than one mutation (Pires and Ascher, 2016).

4. Discussion

The main aims of protein engineering approaches are usually to improve affinity/specificity or to prevent aggregation and increase solubility and stability while not changing affinity/specificity. Although there are some trade-offs during these processes, there are many successful examples in the literature that improved the biophysical characteristics of antibodies.

Enever et al. used a new approach called phage display stress selection to screen for more stable human nanobodies (Enever et al., 2015). Their goals were to improve thermodynamic stability and to make nanobodies resistant to aggregation. They generated error-prone PCR phage libraries and subjected these libraries to various stress conditions. Stress conditions were related to temperature (incubation at 50–80 °C for various amounts of time), pH (incubation at pH 3.2 for various amounts of time), and protease (incubation with trypsin, elastase, leucozyme). Selection results revealed that beneficial mutations (both on CDRs and framework residues) were common to most of the stress conditions. This means that antibodies tend to mutate generic amino acids to improve their biophysical properties.

Dudgeon et al. introduced a general strategy to improve biophysical properties of antibody variable domains (Dudgeon et al., 2012). They identified specific positions in CDR regions (28, 30–33, 35 in V_H and 24, 49–53, 56 in V_L) and mutated those to aspartate or glutamate. This strategy led to increased aggregation resistance, which is advantageous for both diagnostic and therapeutic applications. Although most of those mutations were located in CDR regions, they showed that binding performances were not significantly affected for nearly half of the mutants.

Courtois et al. rationally designed a biobetter drug candidate by mutating or engineering aggregation-prone residues of a Fab fragment (Courtois et al., 2016). They removed aggregation-prone residues by single point mutations (hydrophobic residues to charged aspartate or lysine) and found that stability increased up to 4-fold. They also added a glycosylation site near aggregation-prone regions to increase solubility and up to 3-fold increases in stability were obtained. Most importantly, these engineering approaches did not alter binding to the target.

Before designing mutations to decrease aggregation and/or increase stability of antibodies, three important points should be considered carefully: (i) CDR regions of the sequence should not be selected for mutation although they have high predicted scores because they are usually important for antigen binding and affinity/specificity might be impaired. (ii) Exposed hydrophobic amino acids are widely known to contribute to aggregation, and those residues should be considered first for mutation. They are preferentially mutated to hydrophilic, even charged amino acids such as aspartate, glutamate (Dudgeon et al., 2012), or lysine (Courtois et al., 2016) to circumvent aggregation problems. (iii) Designed mutations should also be compared with a natural repertoire because mutating a residue to its naturally conserved amino acid might improve its properties. The abYsis database is a web-based tool that integrates sequence data from the European Molecular Biology Laboratory European Nucleotide Archive (EMBL-ENA) and structure data from the Protein Data Bank (PDB). The abYsis database can be used to determine location-specific amino acid distribution of the natural repertoires of different organisms (Swindells et al., 2017).

It is important to note that there could be some trade-offs while improving the desired properties of an antibody (solubility, stability, affinity). Thus, the designed change should be considered for all properties. For example, while a mutation increases the solubility, it might also decrease stability at the same time. Affinity maturation can lead to a better binder but this higher affinity antibody might fail in the development phase due to its poor biophysical characteristics. It is important to keep in mind that trade-offs can occur while improving antibody fragments and one should design their computational/experimental setup accordingly.

Acknowledgments

We would like to thank the İzmir Biomedicine and Genome Center and YÖK (Council of Higher Education) 100/2000 fellowship program for funding our research group. We thank Hasan Buğra Çoban for his valuable input during writing process. We thank all of our research group members for carefully reviewing this article before submission.

References

Agrawal NJ, Kumar S, Wang XL, Helk, B, Singh SK, Trout BL (2011). Aggregation in protein-based biotherapeutics: computational studies and tools to identify aggregation-prone Regions. Journal of Pharmaceutical Sciences 100: 5081-5095.

Bain B, Brazil M (2003). Adalimumab. Nature Reviews Drug Discovery 2: 693-694.

Bava KA, Gromiha MM, Uedaira H, Kitajima K, Sarai A (2004). ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic Acids Research 32: D120-D121.

Beerten J, Van Durme J, Gallardo R, Capriotti E, Serpell L, Rousseau F, Schymkowitz J (2015). WALTZ-DB: a benchmark database of amyloidogenic hexapeptides. Bioinformatics 31: 1698-1700.

Bradbury A, Pluckthun A (2015). Standardize antibodies used in research. Nature 518: 27-29.

Brodel AK, Isalan M, Jaramillo A (2018). Engineering of biomolecules by bacteriophage directed evolution. Current Opinion in Biotechnology 51: 32-38.

Capriotti E, Fariselli P, Casadio R (2004). A neural-network-based method for predicting protein stability changes upon single point mutations. Bioinformatics 20: 63-68.

Capriotti E, Fariselli P, Casadio R (2005). I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research 33: W306-W310.

Chan CE, Lim AP, MacAry PA, Hanson BJ (2014). The role of phage display in therapeutic antibody discovery. International Immunology 26: 649-657.

Chao G, Lau W L, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD (2006). Isolating and engineering human antibodies using yeast surface display. Nature Protocols 1: 755-768.

Cheng JL, Randall A, Baldi P (2006). Prediction of protein stability changes for single-site mutations using support vector machines. Proteins-Structure Function and Bioinformatics 62: 1125-1132.

Chiu ML, Gilliland GL (2016). Engineering antibody therapeutics. Current Opinion in Structural Biology 38: 163-173.

Conchillo-Sole O, de Groot NS, Aviles FX, Vendrell J, Daura X, Ventura S (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics 8: 65.

Courtois F, Agrawal NJ, Lauer TM, Trout BL (2016). Rational design of therapeutic mAbs against aggregation through protein engineering and incorporation of glycosylation motifs applied to bevacizumab. MAbs 8: 99-112.

Dubay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendruscolo M (2004). Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. Journal of Molecular Biology 341: 1317-1326.

Dudgeon K, Rouet R, Kokmeijer I, Schofield P, Stolp J, Langley D, Stock D, Christ D (2012). General strategy for the generation of human antibody variable domains with increased aggregation resistance. Proceedings of the National Academy of Sciences of the United States of America 109: 10879-10884.

Elgundi Z, Reslan M, Cruz E, Sifniotis V, Kayser V (2017). The state-of-play and future of antibody therapeutics. Advanced Drug Delivery Reviews 122: 2-19.

Enever C, Pupecka-Swider M, Sepp A (2015). Stress selections on domain antibodies: 'What doesn't kill you makes you stronger'. Protein Engineering Design & Selection 28: 59-66.

Feldhaus MJ, Siegel RW (2004). Yeast display of antibody fragments: a discovery and characterization platform. Journal of Immunological Methods 290: 69-80.

Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nature Biotechnology 22: 1302-1306.

Fink AL (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding & Design 3: R9-R23.

Frenzel A, Kugler J, Helmsing S, Meier D, Schirrmann T, Hust M, Dubel S (2017). Designing human antibodies by phage display. Transfusion Medicine and Hemotherapy 44: 312-318.

García Merino A (2011). Monoclonal antibodies. Basic features. Neurología (English Edition) 26: 301-306.

Gasior P, Kotulska M (2014). FISH Amyloid - a new method for finding amyloidogenic segments in proteins based on site specific co-occurrence of aminoacids. BMC Bioinformatics 15: 54.

Georgiou G, Poetschke HL, Stathopoulos C, Francisco JA (1993). Practical applications of engineering gram-negative bacterial-cell surfaces. Trends in Biotechnology 11: 6-10.

Groves MAT, Amanuel L, Campbell JI, Rees DG, Sridharan S, Finch DK, Lowe DC, Vaughan TJ (2014). Antibody VH and VL recombination using phage and ribosome display technologies reveals distinct structural routes to affinity improvements with VH-VL interface residues providing important structural diversity. MAbs 6: 236-245.

Gustafson HH, Olshefsky A, Sylvestre M, Sellers DL, Pun SH (2018). Current state of in vivo panning technologies: Designing specificity and affinity into the future of drug targeting. Advanced Drug Delivery Reviews 130: 39-49.

Hairul Bahara NH, Tye GJ, Choong YS, Ong EB, Ismail A, Lim TS (2013). Phage display antibodies for diagnostic applications. Biologicals 41: 209-216.

Hanes J, Pluckthun A (1997). In vitro selection and evolution of functional proteins by using ribosome display. Proceedings of the National Academy of Sciences of the United States of America 94: 4937-4942.

Harel Inbar N, Benhar I (2012). Selection of antibodies from synthetic antibody libraries. Archives of Biochemistry and Biophysics 526: 87-98.

Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J (2017). Protein-Sol: a web tool for predicting protein solubility from sequence. Bioinformatics 33: 3098-3100.

Ho M, Nagata S, Pastan I (2006). Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. Proceedings of the National Academy of Sciences of the United States of America 103: 9637-9642.

Ho M, Pastan I (2009). Mammalian cell display for antibody engineering. Methods in Molecular Biology 525: 337-352.

Jeong KJ, Seo MJ, Iverson BL, Georgiou G (2007). APEx 2-hybrid, a quantitative protein-protein interaction assay for antibody discovery and engineering. Proceedings of the National Academy of Sciences of the United States of America 104: 8247-8252.

Jijakli K, Khraiwesh B, Fu W, Luo L, Alzahmi A, Koussa J, Chaiboonchoe A, Kirmizialtin S, Yen L, Salehi-Ashtiani K (2016). The in vitro selection world. Methods 106: 3-13.

Kamtekar S, Schiffer JM, Xiong HY, Babik JM, Hecht MH (1993). Protein design by binary patterning of polar and nonpolar amino-acids. Science 262: 1680-1685.

Kaplon H, Reichert JM (2018). Antibodies to watch in 2018. MAbs 10: 183-203.

Karimi M, Mirshekari H, Moosavi Basri SM, Bahrami S, Moghoofei M, Hamblin MR (2016). Bacteriophages and phage-inspired nanocarriers for targeted delivery of therapeutic cargos. Advanced Drug Delivery Reviews 106: 45-62.

Kiguchi Y, Oyama H, Morita I, Katayama E, Fujita M, Narasaki M, Yokoyama A, Kobayashi N (2018). Antibodies and engineered antibody fragments against M13 filamentous phage to facilitate phage-display-based molecular breeding. Biological & Pharmaceutical Bulletin 41: 1062-1070.

Ledsgaard L, Kilstrup M, Karatt-Vellatt A, McCafferty J, Laustsen AH (2018). Basics of antibody phage display technology. Toxins (Basel) 10: E236.

Li W, Caberoy NB (2010). New perspective for phage display as an e%cient and versatile technology of functional proteomics. Applied Microbiology and Biotechnology 85: 909-919.

Liu R, Li X, Xiao W, Lam KS (2017). Tumor-targeting peptides from combinatorial libraries. Advanced Drug Delivery Reviews 110-111: 13-37.

Loset GA, Sandlie I (2012). Next generation phage display by use of pVII and pIX as display scaffolds. Methods 58: 40-46.

Mei M, Zhou Y, Peng W, Yu C, Ma L, Zhang G, Yi L (2017). Application of modified yeast surface display technologies for non-Antibody protein engineering. Microbiological Research 196: 118-128.

Nagumo Y, Fujiwara K, Horisawa K, Yanagawa H, Doi N (2016). PURE mRNA display for in vitro selection of single-chain antibodies. Journal of Biochemistry 159: 519-526.

Nelson AL (2010). Antibody fragments: hope and hype. MAbs 2: 77-83.

Nixon AE, Sexton DJ, Ladner RC (2014). Drugs derived from phage display: from candidate identification to clinical practice. MAbs 6: 73-85.

Paladin L, Piovesan D, Tosatto SCE (2017). SODA: prediction of protein solubility from disorder and aggregation propensity. Nucleic Acids Research 45: W236-W240.

Pandurangan AP, Ochoa-Montano B, Ascher DB, Blundell TL (2017). SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Research 45: W229-W235.

Parthiban V, Gromiha MM, Schomburg D (2006). CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Research 34: W239-W242.

Pawar AP, DuBay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM (2005). Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. Journal of Molecular Biology 350: 379-392.

Pires DEV, Ascher DB (2016). mCSM-AB: a web server for predicting antibody-antigen a%nity changes upon mutation with graph-based signatures. Nucleic Acids Research 44: W469-W473.

Qin CF, Li GC (2014). Mammalian cell display technology coupling with AID induced SHM in vitro: an ideal approach to the production of therapeutic antibodies. International Immunopharmacology 23: 380-386.

Sergeeva A, Kolonin MG, Molldrem JJ, Pasqualini R, Arap W (2006). Display technologies: application for the discovery of drug and gene delivery agents. Advanced Drug Delivery Reviews 58: 1622-1654.

Smith GP (1985). Filamentous fusion phage - novel expression vectors that display cloned antigens on the virion surface. Science 228: 1315-1317.

Sormanni P, Amery L, Ekizoglou S, Vendruscolo M, Popovic B (2017). Rapid and accurate in silico solubility screening of a monoclonal antibody library. Scientific Reports 7: 8200.

Sormanni P, Aprile FA, Vendruscolo M (2015). The CamSol method of rational design of protein mutants with enhanced solubility. Journal of Molecular Biology 427: 478-490.

Stohl W, Hilbert DM (2012). The discovery and development of belimumab: the anti-BLyS-lupus connection. Nature Biotechnology 30: 69-77.

Swindells MB, Porter CT, Couch M, Hurst J, Abhinandan KR, Nielsen JH, Macindoe G, Hetherington J, Martin ACR (2017). abYsis: Integrated antibody sequence and structure-management, analysis, and prediction. Journal of Molecular Biology 429: 356-364.

Tartaglia GG, Vendruscolo M (2008). The Zyggregator method for predicting protein aggregation propensities. Chemical Society Reviews 37: 1395-1401.

Teixeira D, Gonzalez-Pajuelo, M (2018). Phage display technology for selection of antibody fragments. In: Sarmento B, Das Neves J (editors). Biomedical Applications of Functionalized Nanomaterials. Amsterdam, the Netherlands: Elsevier, pp. 67-88.

Thiltgen G, Goldstein RA (2012). Assessing predictors of changes in protein stability upon mutation using self-consistency. PLoS One 7: e46084.

van der Kant R, van der Kant R, Karow-Zwick AR, Durme JV, Blech M, Gallardo R, Seeliger D, Aßfalg K, Baatsen P, Compernolle G et al. (2017). Prediction and reduction of the aggregation of monoclonal antibodies. Journal of Molecular Biology 429: 1244-1261.

Wainreb G, Wolf L, Ashkenazy H, Dehouck Y, Ben-Tal N (2011). Protein stability: a single recorded mutation aids in predicting the effects of other mutations in the same amino acid site. Bioinformatics 27: 3286-3292.

Walsh I, Seno F, Tosatto SCE, Trovato A (2014). PASTA 2.0: an improved server for protein aggregation prediction. Nucleic Acids Research 42: W301-W307.

Wang XL, Singh SK, Kumar S (2010). Potential aggregation-prone regions in complementarity-determining regions of antibodies and their contribution towards antigen recognition: a computational analysis. Pharmaceutical Research 27: 1512-1529.

West MW, Wang WX, Patterson J, Mancias J D, Beasley JR, Hecht MH (1999). De novo amyloid proteins from designed combinatorial libraries. Proceedings of the National Academy of Sciences of the United States of America 96: 11211-11216.

Worn A, Pluckthun A (1999). Different equilibrium stability behavior of ScFv fragments: identification, classification, and improvement by protein engineering. Biochemistry 38: 8739-8750.

Worn A, Pluckthun A (2001). Stability engineering of antibody single-chain Fv fragments. Journal of Molecular Biology 305: 989-1010.

Xiao XD, Chen Y, Mugabe S, Gao C, Tkaczyk C, Mazor Y, Pavlik P, Wu H, Dall'Acqua W, Chowdhury PS (2017). A high-throughput platform for population reformatting and mammalian expression of phage display libraries to enable functional screening as full-length IgG. MAbs 9: 996-1006.

Yageta S, Lauer TM, Trout BL, Honda S (2015). Conformational and colloidal stabilities of isolated constant domains of human immunoglobulin G and their impact on antibody aggregation under acidic conditions. Molecular Pharmaceutics 12: 1443-1455.

Yang Y, Niroula A, Shen BR, Vihinen M (2016). PON-Sol: prediction of effects of amino acid substitutions on protein solubility. Bioinformatics 32: 2032-2034.

Zambrano R, Jamroz M, Szczasiuk A, Pujols J, Kmiecik S, Ventura S (2015). AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures. Nucleic Acids Research 43: W306-W313.

Received: 18 December 2019 | Revised: 23 March 2020 | Accepted: 6 June 2020

DOI: 10.1002/prot.25964

RESEARCH ARTICLE

PROTEINS WILEY

Conformational changes in a Vernier zone region: Implications for antibody dual specificity

Merve Arslan[1,2] | Dilara Karadag[1] | Sibel Kalyoncu[1]

[1]Izmir Biomedicine and Genome Center, Izmir, Turkey

[2]Izmir Biomedicine and Genome Institute, Dokuz Eylül University, Izmir, Turkey

Correspondence
Sibel Kalyoncu, Izmir Biomedicine and Genome Center, Izmir, Turkey.
Email: sibel.kalyoncu@ibg.edu.tr

Funding information
Türkiye Bilimsel ve Teknolojik Arastirma Kurumu, Grant/Award Number: 119Z161; Yükseköğretim Kurulu, Grant/Award Number: 100/2000

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1002/prot.25964.

Abstract

Understanding the determinants of antibody specificity is one of the challenging tasks in antibody development. Monospecific antibodies are still dominant in approved antibody therapeutics but there is a significant body of work to show that multispecific antibodies can increase the overall therapeutic effect. Dual-specific or "Two-in-One" antibodies can bind to two different antigens separately with the same antigen-binding site as opposed to bispecifics, which simultaneously bind to two dif- ferent antigens through separate antigen-binding units. These nonstandard dual- specific antibodies were recently shown to be promising for new antibody-based therapeutics. Here, we physicochemically and structurally analyzed six different anti- bodies of which two are monospecific and four are dual-specific antibodies derived from monospecific templates to gain insight about dual-specificity determinants. These dual-specific antibodies can target both human epidermal growth factor recep- tor 2 and vascular endothelial growth factor at different binding affinities. We showed that a particular region of clustered Vernier zone residues might play key roles in gaining dual specificity. While there are minimal intramolecular interactions between a certain Vernier zone region, namely LV4 and LCDR1 of monospecific tem- plate, there is a significant structural change and consequently close contact forma- tion between LV4-LCDR1 loops of derived dual-specific antibodies. Although Vernier zone residues were previously shown to be important for humanization applications, they are mostly underestimated in the literature. Here, we also aim to resurrect Ver- nier zone residues for antibody engineering efforts.

KEYWORDS

antibody, dual specific, human epidermal growth factor receptor 2 (HER2), specificity, vascular endothelial growth factor (VEGF), Vernier zone

1 | INTRODUCTION

Monoclonal antibodies are one of the most important biological drugs being developed for targeted therapy. The specificity of antibody- antigen interactions is one of the main parameters for the success of antibodies for therapeutic purposes. Multi-specificity (or cross-reac- tivity) of antibodies is an important phenomenon due to their roles in immune recognition.[1] Although multi-specificity is often associated with self-reactivity and autoimmunity, it turned out to be a conserved feature of the immune system.[2] The ability of antibodies to become multi-specific with somatic mutations affects the prevalence of them in antibody repertoire.[3,4] Natural antibodies are not exclusively spe- cific and some of them are known to interact with more than one anti- gen with decent affinities.[5] It is recently shown that multi-specific antibodies could be an important feature of the immune system to enhance its repertoire.[6,7] Although multi-specificity might lead to

autoimmunity, there is a fine balance between attacking pathogens and removing autoreactivity as a result of evolutionary pressure.[8,9]

Due to genetic heterogeneity and complex biological pathways, cancer and certain infectious diseases are generally difficult to treat with a monospecific therapeutic agent. In recent years, different types of antibodies such as dual-specific and bispecific antibodies have been developed to increase the therapeutic effect.[10] Bispecific antibodies can simultaneously bind to two different antigens.[11] There are currently three bispecific antibodies among 87 antibodies approved either in EU or in US, and this number is expected to increase due to their clinic successes.[12] Although there are many formats of bispecific antibodies, the most common form is having two different antibodies at separate arms each targeting different antigens.[13] On the other hand, dual specific, also named "Two-in-One" antibodies, can bind to two different antigens separately with the same antigen-binding site.[14] Modulation of antibody specificity can lead to many effective biopharmaceutical and diagnostic applications. Antibodies mostly uti- lize heavy chain complementarity determining regions (CDRs) as the main antigen-binding determinants. Thus, light chain CDRs can be available to engineer for affinity to a second antigen. Some antibody engineering techniques have been successfully applied to generate dual-specific antibodies from a monospecific antibody.[14-20] Bostrom et al used a phage display library to derive novel dual-specific anti- bodies from a monospecific anti-Human Epidermal Growth Factor Receptor 2 (HER2) antibody.[20] It was shown that several limited mutations on light chain CDRs are sufficient to obtain dual-specific antibodies recognizing both HER2 and vascular endothelial growth factor (VEGF). This pioneering approach proved that monospecific antibodies can be designed to develop dual (or multi) specificity through mutations in the light chain CDR regions.

Antibody specificity is an important and complex issue in anti- body development, and it is still not completely understood.[21] The relationship between mono-/multi-/non-specificity is very hard to understand. Although multi-specificity could be important for immune repertoire and biopharmaceuticals, caution must be taken not to impair antibody developability.[22] It was shown that approved antibodies are more specific than those in clinical trials[18] and this specific- ity difference might depend on many factors such as the aliphatic content of CDRs.[23] Several properties such as hydrophobicity, iso- electric point, glycosylation, and charge are reported to affect anti- body specificity.[24-27]

Antibody specificity and/or affinity modulation efforts are mostly based on CDR regions. The main differences between all antibody vari- able domains are content, structure, and conformations of CDR loops. Non-CDR framework regions are mostly conserved and have a high degree of structural conservation forming a core β-sheet structure.[28] Although several studies point out that non-CDR regions are important for biophysical properties[29] and humanization,[30,31] their roles on affin- ity/specificity are underestimated.[29,32] One of the important non-CDR regions is the Vernier zone. Vernier zone residues are located in the framework regions and underlie the complementary determining regions (CDRs). These residues potentially affect the conformation of CDR loop structures.[33,34] Antibody humanization approaches mostly utilize Vernier zone residues to reshape CDR loops.[30,35] Back mutations on the Vernier zone can provide the desired canonical structure of CDRs to obtain restored binding affinity.[36] However, there is no study investigating the relationship between Vernier zone residues and anti- body specificity.[28] Although Vernier zone residues are mostly engineered for humanization efforts to regain/improve affinity, it can be hypothesized that features of Vernier zone residues might also affect many antibody properties such as binding specificity.

In this study, we found that one of the Vernier zone regions play important roles in gaining dual specificity from a monospecific antibody. When sequence and structure of parental monospecific anti-HER2 antibody are compared with those of dual-specific anti-VEGF/HER2 and monospecific anti-VEGF antibody variants, one of Vernier zone regions interacting with engineered LCDR1 stands out. The canonical structure of LCDR1 drastically changes and makes close contacts with the partic- ular Vernier zone region in dual-specific antibodies. This study shows that previously underestimated Vernier zone regions might help us to modulate antibody specificity in a controlled manner.

2 | MATERIALS AND METHODS

2.1 | Homology modeling

Data for all antibody sequences and properties were collected from Bostrom et al.[20] ROSIE antibody server[37] (Rosetta Online Server[38]) was used to build homology models of 3-1, bH3, bH4, and bH1-81. Because ROSIE antibody server models only variable regions of anti- bodies, sequences of heavy and light variable regions were given as input. Homolog templates were chosen based on BLAST from anti- body crystal structures in the Protein Data Bank (PDB). Template sea- rch was made independently for each CDR and framework regions, additional remodeling was done for the HCDR3 loop. Lowest energy refined models of 3-1, bH3, bH4, bH1-81, and crystal structures of Herceptin and bH1 were used for further steps. 3D protein structures of herceptin-HER2 (PDB ID: 1N8Z), bH1-VEGF (PDB ID: 3BDY), and bH1-HER2 (PDB ID: 3BE1) complexes were extracted from the PDB.

2.2 | Interface refinement

In order to obtain structures of 3-1, bH3, bH4, and bH1-81 antibodies in complex with VEGF and/or HER2, the HADDOCK-Refinement interface program was used.[39] HADDOCK-Refinement is a molecular dynamics simulated refinement module under HADDOCK 2.2 web server.[40] Refined homology models of 3-1, bH3, bH4, bH1-81 and their corresponding antigens (VEGF or HER2) were given as input. The same interface as in Herceptin and bH1 complex structures was used. Water refinement was performed with a default set of parame- ters. Structures with the best HADDOCK score were chosen for fur- ther analysis. HADDOCK score was calculated by Equation (1) below and parameters for data quality were represented in Supplementary Figures 1-7 and Table 1. The score is calculated as:

	LCDR1										LCDR2				LCDR3			
Interaction	28	29	30	30a	30b	30c	30d	31	32	33	50	51	52	53	91	92	93	94
Herceptin-HER2	D	V	N	-	-	-	-	T	A	V	S	A	S	F	H	Y	T	T
			LV4							LV4	LV3	LV3 LV4	LV3 LV4	LV3		Bind		Bind
bH1-HER2			Bind															
	D	I	P	R	S	I	S	G	Y	V	W	G	S	Y	H	Y	T	T
bH3-HER2						LV1 LV4	LV4		Bind		LV3	LV3 LV4	LV4	Bind	Bind	Bind		Bind
	D	I	G	L	-	-	-	G	S	V	W	A	S	Y	H	Y	T	T
bH4-HER2																		
	LV4	LV4						LV4				LV3 LV4	LV4	LV3	LV2	Bind	LV1	Bind
	D	I	R	S	-	-	-	G	S	V	W	G	S	Y	H	Y	T	T
bH1-81-HER2		LV4									LV3	LV3 LV4	LV3 LV4	LV3	LV2			Bind
															Bind			

N	I	A	K	T	I	S	G	Y	V	W	G	S	F	H	Y	S	S
	LV2 LV4	LV4		Bind	LV4	LV4 Bind	LV4	LV4		LV3	LV2 LV3	LV3 LV4	LV3				

TAB L E 1 Variable light chain paratope interactions of antibodies with HER2 antigen [Color table can be viewed at wileyonlinelibrary.com]

Note: LV: light chain Vernier zone region. Blue-colored box represents intramolecular interactions (<3.5 Å) of represented light chain residue with at least one Vernier zone region residue (refer to Figure 2 for LV1-5 regions). Yellow-colored box represents intermolecular H-bond and/or salt bridge interactions of represented light chain residue with HER2 epitope residues.

$$\text{HADDOCK score} = 1{:}0 \times E_{vdw} + 0{:}2 \times E_{elec} + 1{:}0 \times E_{desol} + 0{:}1 \times E_{air}$$

where E_{vdw} is the intermolecular van der Waals energy, E_{elec} is the intermolecular electrostatic energy, E_{desol} is the desolvation energy, and E_{air} is the ambiguous interaction restraints energy.[41]

2.3 | Data analysis

Molecular visualization of structures was done by PyMOL (The PyMOL Molecular Graphics System, Version 2.2 Schrödinger, LLC). Hydrogen bonds (H-bonds), salt bridges, and hydrophobic interactions within interfacing residues of antibody-antigen complexes were defined by PDBePISA-Interfaces tool[42] and further confirmed by PyMOL 2.2 software. Buried surface area (BSA) was calculated from the ratio of BSA to accessible surface area. Amino acid sequences of

V_H and V_L domains of all antibodies were aligned with Clustal Omega.[43] CDR and Vernier zone sequences[44] of V_H and V_L domains were determined according to the Kabat numbering scheme for fur- ther analysis.

The number and type of intermolecular interactions were also analyzed by using PRODIGY (PROtein binDing enerGY prediction) tool, which is a webserver to predict the binding affinity of protein- protein complexes from their 3D structures based on intermolecular contacts and properties derived from the non-interface surface.[45] Default intermolecular contacts distance cutoff of 5.5 Å and default temperature of 25ºC[46] were used for the analysis. The total number of interfacial contacts are the sum of reported charged-charged, charged-polar, charged-apolar, polar-polar, polar-apolar, and apolar- apolar contacts between particular antibody-antigen intermolecular interactions.

2.4 | Computational alanine scanning

Roles of light chain Vernier zone residues on antibody stability and affinity were analyzed by using structure/sequence-based tools, which measures the effects of alanine mutations on certain positions. Generated models or sequences of 3-1, bH3, bH4, bH1-81 and crystal structures or sequences of Herceptin-HER2 (1N8Z), bH1-VEGF (3BDY), and bH1-HER2 (3BE1) were used as input for computational alanine scanning. Three tools based on structural information were used: Cutoff Scanning Matrix[47] (mCSM, http://biosig.unimelb.edu.au/ mcsm/), Site-Directed Mutator[48] (SDM, http://marid.bioc.cam.ac.uk/ sdm2), and mCSM-AB[49](http://biosig.unimelb.edu.au/mcsm_ab/). Two tools based on sequence information were used: I-Mutant2.0[50]
(http://folding.biofold.org/cgi-bin/i-mutant2.0.cgi) and MUPro[51] (https://www.ics.uci.edu/~baldig/mutation.html). Default tempera- ture and pH of 25ºC and 7 were used in I-Mutant2.0. Corresponding Vernier zone residues were substituted into alanine to measure the predicted change (WT/alanine) in Gibbs free energy (.6.6*G*, kcal/mol). mCSM-AB is the only tool predicting antibody-antigen affinity changes upon mutation, rest gives stabilizing/destabilizing predictions.

3 | RESULTS AND DISCUSSION

Antibody sequences of dual-specific antibodies are obtained from the study of Bostrom et al.[3] The aim of their study was to generate a "Two-in-One" dual-specific antibody that would bind to both HER2 and VEGF with reasonable affinities. First, they designed a phage library based on diversifying 12 positions on light chain CDRs of HER2 binding Herceptin (residues 28-32 for LCDR1, 50, 51, 53 for
LCDR2, 91-94 for LCDR3). Loop length was also varied by inserting 0-5 residues on LCDR1 and 0-2 residues on LCDR3. Then, monospe- cific and dual-specific clones were isolated and analyzed for HER2 and VEGF binding. On the following study, they found that interac- tions of HER2/VEGF dual-specific antibodies are entropy-driven while parent herceptin-HER2 interaction is mostly enthalpy-driven.[2] Because dual-specific antibodies have very high potential for the development of next-generation therapeutics, further studies to explore this phenomenon are definitely needed.

In this study, we analyzed structural and physicochemical characteristics of monospecific and dual-specific antibodies for HER2/VEGF dual binding to gain insights about the antibody specificity phenomenon. There are crystal structures available for monospecific Herceptin (in complex with HER2, PDB ID:1N8Z) and dual-specific bH1 (in complex with HER2, PDB ID:3BE1 and VEGF, PDB ID:3BDY, separately). Struc- tures of the rest of the antibodies 3-1 (VEGF monospecific), bH3, bH4, bH1-81 (HER2/VEGF dual specific) were homology modeled and their interaction surfaces were refined with structures of their respective anti- gens. Structure and interaction analysis was performed in detail and some interesting clues about antibody dual specificity were obtained.

3.1 | Dual specificity is mediated by light chain

HER2 binding affinities of all antibodies are in the low nanomolar range of K_ds with the highest binding affinity for Herceptin, as expected (Figure 1A). However, VEGF binding affinities vary drasti- cally. Monospecific 3-1 and dual-specific bH1-81 have the highest VEGF binding affinity with K_ds of 15 and 41 nM, respectively. Both bH3 and bH4 have the worst VEGF binding with K_ds in the micromo- lar range. It is important to note that the heavy chain sequence of all listed antibodies is the same. As expected, when the number of inter- facial contacts and interface area of antibody paratopes are compared, there is definitely a shift to light chain interaction for VEGF binding (Figure 1). While Herceptin has almost the same interface areas for its heavy and light chains, all VEGF binding interactions show a signifi- cantly larger light chain interface area (Figure 1A). This obvious shift is also confirmed with the increasing number of light chain interfacial contacts in VEGF binding except for bH3 and bH4, which are the worst VEGF binders (Figure 1B). When V_H and V_L structures of all antibodies are overlayed, HCDR loops align very well but there is a clear deviation in LCDR loops especially LCDR1 (Figure 1C). This shows that the light chain plays a key role in VEGF affinity, conse- quently dual specificity. Interestingly, dual-specific bH1-81 also has a larger light chain interface area (Figure 1A) and a higher number of

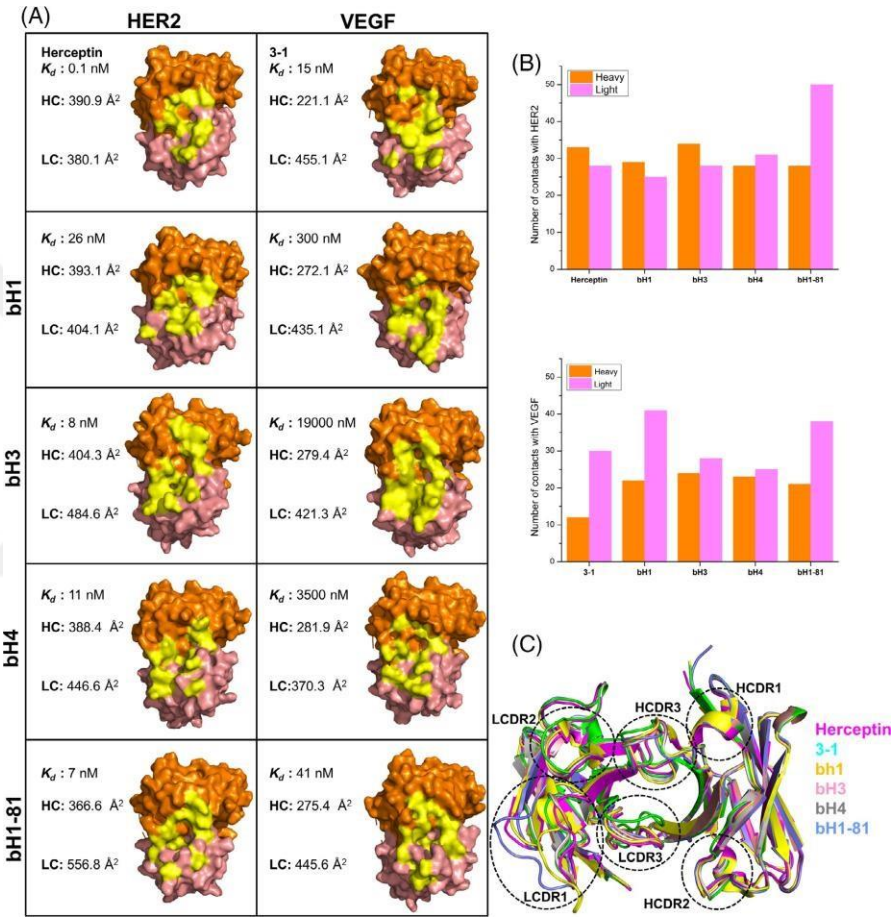
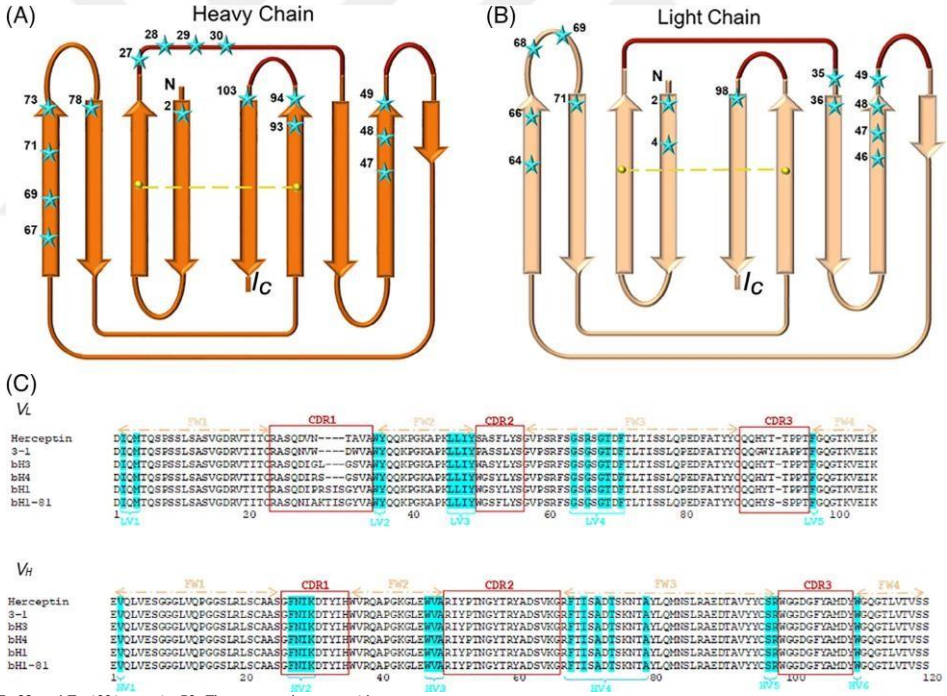


FIG U R E 1 Paratope interaction interfaces of all antibodies used in this study. A, Interfaces of monospecific Herceptin and 3-1 are on the top, those of dual-specific antibodies for corresponding antigen [human epidermal growth factor receptor 2 (HER2): left side, vascular endothelial growth factor (VEGF): right side] are followed. Surfaces of heavy, light, and interface regions are represented in pink, orange, and yellow, respectively. K_d: dissociation constant, LC: interface area for light chain, HC: interface area for the heavy chain. B, Number of interfacial contacts

(charged-charged, charged-polar, charged-apolar, polar-polar, polar-apolar, and apolar-apolar) for HER2 (top) and VEGF (bottom) bindings. C, Structure overlay of variable regions of Herceptin (magenta), 3-1 (cyan), bH1 (yellow), bH3 (pink), bH4 (gray), and bH1-81 (blue)

contacts with HER2 when its heavy chain interaction is almost the same as Herceptin's (Figure 1B); however, this trend does not increase its affinity for HER2. This also confirms that HER2 binding is mostly mediated by a heavy chain, and dual specificity for VEGF is mostly mediated by the light chain.

There are many common core residues on the heavy chain for HER2 binding interaction, which are Arg50, Arg59, Gly99, and Tyr100A (Supplementary Tables 2-6). While Arg50-Glu558, Gly99-Lys593, and Tyr100A-Pro571 interactions are common for all HER2 binding interactions, Arg59 changes its binding partner and interaction type in dual-specific antibodies (Asp560 for Herceptin, Gln561/Glu558 for dual specifics). Losing Arg50-Asp560 salt bridge in dual-specific antibodies might have a role in a 70-260x fold affinity decrease in HER2 affinity (Figure 1A). There are also some common core residues on the heavy chain for VEGF binding interaction, which are Arg50 and Gly99 (Supplementary Tables 7-11). Both Arg50-His86 and Gly99-Gln89



Tyr33 and Tyr100A, not Arg50. There are only two residues on VEGF, His86, and Gln89, which form contact with heavy chain compared to at least six residues for HER2. This is in accordance with having a larger contact area of light chain for VEGF binding. Therefore, light chain contacts and interactions are comprehensively analyzed to get an insight about dual specificity.

3.2 | Vernier zone residues are clustered as separate regions

Vernier zone residues are known to be important for CDR conforma- tion through Vernier zone-CDR interactions. They are usually used in humanization approaches as back mutation candidates to restore binding affinity of CDR-grafted humanized antibodies.[52] Although there is a Kabat numbering based assignment of Vernier zone residues,[33,44] there is no comprehensive analysis of these regions.

Here, we clustered Vernier zone residues around stems of CDR regions; we named and numbered them accordingly (Figure 2). There are six and five Vernier zone regions for heavy (HV1-6) and light (LV1-5) chains, respectively (Figure 2C). Unlike the rest of other regions, LV4 and HV4 are not subsequent or precedent of any CDR loop sequence. They are part of a loop facing antigen in the same direction as other CDRs. These loops reside on Framework-3 and they are non-hypervariable.[53,54] Some studies called this loop "CDR4" due to its potential effect on antigen binding.[55] In this study, we show that one of Vernier zone regions, LV4, makes an extraordinary close con- tact with LCDR1 in dual-specific antibodies.

3.3 | LCDR1-LV4 interaction is essential for dual specificity

While the number of binding residues on LCDR2 and LCDR3 does not significantly change for VEGF vs HER2, there are more binding

FIG U R E 2 Schematic representation of variable regions of (A) heavy and (B) light chains. Complementarity determining regions (CDRs) are colored in red, Vernier zone residues are labeled as stars in cyan, disulfide bonds are represented in yellow. C, Variable chain sequences of all antibodies used in this study. CDR, framework (FW), and Vernier zone (HV for the heavy chain, LV for the light chain) regions are labeled and numbered according [Color figure can be viewed at wileyonlinelibrary.com]

	LCDR1										LCDR2				LCDR3				
Interaction	28	29	30	30a	30b	30c	30d	31	32	33	50	51	52	53	91	92	93	93a	94
3-1-VEGF	N	V	W	-	-	-	-	D	W	V	P	A	S	S	G	W	Y	I	A
bH1-VEGF	LV4	LV4	LV4					Bind				LV3	LV3	LV3	Bind		Bind	HV3	
	D	I	P	R	S	I	S	G	Y	V	W	G	S	Y	H	Y	T	-	T
bH3-VEGF	LV4	LV4	LV4	Bind	LV4	LV4		LV4	LV1	LV3	LV3 LV4	LV3 LV4	LV3	Bind	Bind	Bind		Bind	
							Bind				Bind								
bH4-VEGF	D	I	G	L	-	-	-	G	S	V	W	A	S	Y	H	Y	T	-	T
		LV4	Bind	LV4				LV4	LV4		LV3	LV2 LV4	LV3 LV4	LV3	LV2		LV1		Bind
															Bind				
	D	I	R	S	-	-	-	G	S	V	W	G	S	Y	H	Y	T	-	T
LV4	LV4	Bind	LV4				LV4	Bind			LV2 LV3 LV4	LV3 LV4	LV3	Bind		LV1		Bind	

bH1-81-VEGF			Bind															
N	I	A	K	T	I	S	G	Y	V	W	G	S	*F*	H	Y	S	-	S
	LV4	LV4	LV4	Bind	LV4	LV4	LV4	LV4	LV4		LV2 LV3 LV4	LV3 LV4	LV3	Bind	Bind	LV1		Bind
						Bind												

TAB L E 2 Variable light chain paratope interactions of antibodies with VEGF antigen [Color table can be viewed at wileyonlinelibrary.com]

Note: LV: light chain Vernier zone region. Blue-colored box represents intramolecular interactions (<3.5 Å) of represented light chain residue with at least one Vernier zone region residue (refer to Figure 2 for LV1-5 regions). Yellow-colored box represents intermolecular H-bond and/or salt bridge interactions of represented light chain residue with VEGF epitope residues.

residues on LCDR1 for VEGF interaction (Tables 1 and 2). This is in accordance with our previous statement of having more contacts on dual specifics' light chains, however, it seems like this increase is mostly mediated by LCDR1. We also investigated other intramolecular interactions of LCDR residues besides antigen and it is found that there is a significant amount of new contacts formed between LCDRs and Vernier zone residues. We observed that each LCDR region makes intramolecular interactions with specific Vernier zone regions: LCDR1 with LV4 only, LCDR2 with mainly LV3, LCDR3 with LV1 or LV2 (Tables 1 and 2). While there is no significant change in those intramolecular contact maps of LCDR2 and LCDR3 in HER2 vs VEGF binding modes, there is a significant increase in LCDR1-LV4 contact number for VEGF binding. For example, LCDR1s of bH1 and bH4 form more than three times more LV4 interactions for VEGF binding. At first, an increase in LV4 contact number can be attributed to an increase in the length of bH1 and bH1-81 LCDR1s. However, this LCDR1-LV4 interaction increase is true for all antibodies even the ones with the same length of LCDR1 as Herceptin's.

In the study of Bostrom et al,[20] an alanine scanning study was made to bH1 and bH1-44 (derived from bH1 via both heavy and light chain mutations) and LDCR1 residues I29 and Y32 showed a signifi- cant decrease on VEGF binding while keeping HER2 binding the same. I29A/Y32A converted a dual-specific antibody to essentially mono- specific. This confirms our findings for the importance of LCDR1-LV4 interaction because both I29 and Y32 form <3.5 Å interface with the LV4 region (Table 2).

3.4 | Loops of LCDR1 and LV4 come closer for VEGF binding

When the interaction of LCDR1-LV4 is investigated in detail, it is seen that those loops get significantly closer when compared to Herceptin structure (Figure 3A). Not only LV4 loop changes its conformation but also LCDR1 loops change its canonical structure bending toward LV4. This also increases the available surface area of the LCDR1 loop, which correlates with higher LCDR1 contacts formed with both anti- gens (Figure 1). The most drastic increase in LCDR1 contact is for bH1-81-VEGF interaction (Supplementary Tables 8 and 11). This might be due to unusual LCDR1/Ser34 interaction with LV4/Gly68 (Figure 3B). Ser34-Gly68 and Ile29-Gly68 interactions interlock the LCDR1 loop in a conformation that residues in the middle (Thr30B, Ile30C, Ser30D, and Gly31) readily become available for VEGF binding (Supplementary Table 11).

When conformations of all LCDR loops of antibodies are com- pared with Herceptin LCDRs, change in RMSD increases the most for LCDR1 (about 2× more RMSD difference from LCDR2,3, Figure 3A). Most of the LCDR1-LV4 interactions are backbone mediated H-bond contacts (Figure 3B). While LCDR1 and LV4 loops of Herceptin is not close enough for any contact (more than 5 Å distance), all other dual- specific antibodies form close contacts. 3-1 is not dual specific (only binds to VEGF), this might show that forming LCDR1-LV4 interaction is vital for light chain dominated binding of VEGF. There is a signifi- cant LV4-LCDR1 intramolecular interaction increase in VEGF bound

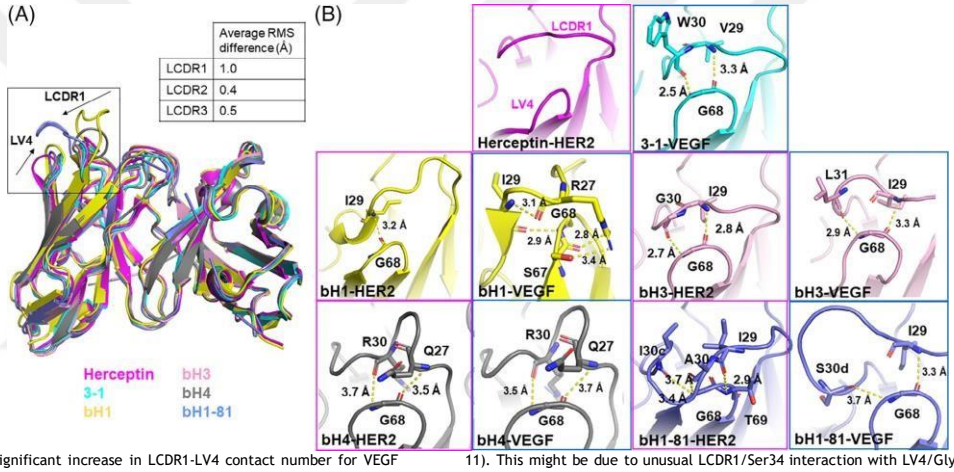


FIG U R E 3 Interaction between LCDR1 and LV4 regions for dual-specific antibodies. A, Variable region structure overlay of Herceptin [human epidermal growth factor receptor 2 (HER2) bound], 3-1 [vascular endothelial growth factor (VEGF) bound], bH1 (HER2 bound), bH3 (HER2 bound), bH4 (HER2 bound), and bH1-81 (HER2 bound). LCDR1-LV4 contact formation is boxed, average root-mean-square deviation (RMSD) values of LCDR loops with respect to those of Herceptin are reported. B, Residue interactions between LCDR1-LV4 loops for VEGF and/or HER2 bound Herceptin, 3-1, bH1, bH3, bH4, and bH1-81. Herceptin does not have any interaction while others have various backbone interactions [Color figure can be viewed at wileyonlinelibrary.com]

form for bH1 compared to that of HER2 bound (Figure 3B). Given that bH1-VEGF and bH1-HER2 are experimentally determined structures, this drastic increase in LCDR-LV4 contacts in VEGF binding strengthens our overall findings of the importance of LCDR1-LV4 interaction for VEGF binding.

Change in the LCDR1 loop structure is expected due to its con- tent and length differences. However, change in the LV4 loop struc- ture is interesting because the content of this Vernier zone is the same for all antibodies and there are no antigen-binding interactions. Both bH1 and bH1-81 have longer LCDR1 loops than other dual spe- cifics, this probably leads to a higher number of LCDR1-VEGF and LCDR1-LV4 interactions. Only bH1-81 has extra binding interactions on LCDR3, this might explain the highest affinity of bH1-81 for VEGF binding compared to other dual specifics (K_d: 41 nM). Also, it was pre- viously shown that longer and flexible CDR loops lead to a decrease in antibody specificity.[56] Both the length and content of LCDR loops could be factors to generate dual-specific antibodies.

The backbone of G68 on LV4 form contacts with LCDR1 in all VEGF-binding antibodies (Figure 3B). Although binding partner of G68 changes based on the length/content of LCDR1, it is mostly around residue number 29 on LCDR1. This might show that G68 is one of the most important Vernier zone residues on LV4 to form con- tact with neighboring LCDR1. Although glycine is really conserved at position 68 (98% for human),[57] mutating it into a hydrophilic residue with longer side chain might make LCDR1-LV4 come closer to modu- late antibody specificity. Changing the canonical structure of LCDR1 through Vernier zone residues and/or LCDR1 engineering might be a new route to search for novel dual-specific antibodies.

3.5 | Dual specifics have more positively charged LCDR1

We also analyzed the physicochemical content of all LCDRs because we previously showed that physicochemical content is key for poly- specificity patterns of therapeutic antibodies.[23] There was no signifi- cant change except total charge on LCDR1 (Supplementary Table 12). While monospecific antibodies (Herceptin and 3-1) have neutral LCDRs, all dual-specific antibodies had positively charged LCDRs except bH3, which is the worst binder of VEGF. The best dual-specific binder of VEGF, bH1-81, has the most positively charged LCDR1 with a +2 charge. It is known that a high frequency of positively charged amino acids on certain CDR loops have been linked to low specificity and increased nucleic acid binding of the antibodies.[4,27,58] Also, it was shown that paratope regions of antibodies are less positively charged compared to other protein-protein interaction interfaces.[59] Although there is a thin line between being dual specific and having low speci- ficity, positively charge amino acids might be favored during the directed evolution of monospecific antibodies into dual specificity.

Determinants of specificity of protein-protein interaction are still challenging to understand.[60] This could be even more challenging for antibody-antigen interactions because they significantly differ from standard protein-protein interactions.[61] In this study, we showed a significant role of an intramolecular interaction involving one of Ver- nier zone regions for gaining dual specificity. Among Vernier zone regions, there are two non-hypervariable antigen-facing loops: one in heavy (HV4) and one in light chain (LV4). We found that LV4 makes an unexpected close contact with LCDR1, this LV4-LCDR1 contact might be important for specificity modulation. In order to see whether LV4 residues could be mutated for antibody specificity engineering efforts, a computational alanine scanning was conducted on light chain Vernier zone regions (Supplementary Figure 8A). Among five regions (LV1-5), LV4 residues stand out by having the least destabilizing effects upon alanine mutations. This is especially true for Gly68, which is the main contact residue for LV4-LCDR1 interaction (Figure 3B). When we analyzed the abYsis database for amino acid frequency distributions of Vernier zone residues,[57] there was high conservation for most of Vernier zone positions and LV4 residues were not that different (Supplementary Figure 8B). However, this does not mean that Vernier zone residues cannot be engineered to improve antibody characteristics because Vernier zone residues are heavily engineered in humanization efforts.[30,31] Besides mutation on Vernier zone residues, length of LV4 and HV4 can also be changed to modulate specificity of antibodies because those loops are the only ones with extendable positions (eg, 66A, 66B for LV4 according to Kabat numbering, Supplementary Figure 8B).

Our main finding was on the light chain probably due to directed evolution efforts made on light chain CDRs for these particular dual- specific antibodies.[20] However, the HV4 region on a heavy chain could be another candidate for such specificity modulation. Some studies called HV4 and LV4 loops "CDR4s" due to their potential effects on antigen binding,[55] our study supports this claim by showing the potential of LV4 for antigen affinity/specificity modulation. Although Vernier zone residues were previously reported to be impor- tant for humanization efforts and binding affinity,[33,62] they are mostly underrepresented in the literature.[28] Our study shows the importance of Vernier zone residues for antibody specificity modulation and it also aims to resurrect Vernier zone regions for antibody engineering efforts.

ACKNOWLEDGMENTS

We would like to thank Izmir Biomedicine and Genome Center (Start- up grant), TUBITAK (Project number: 119Z161) and YOK (Council of Higher Education) 100/2000 fellowship program for funding our research group. We thank Dr. Hasan Bugvra Çoban for his input at the start of this project. We thank all of our research group members for carefully reviewing this article before submission.

ORCID

Sibel Kalyoncu https://orcid.org/0000-0003-2264-0757

REFERENCES

1. Eisen HN. Specificity and degeneracy in antigen recognition: yin and yang in the immune system. *Annu Rev Immunol*. 2001;19:1-21.
2. Cohn M. Degeneracy, mimicry and crossreactivity in immune recogni- tion. *Mol Immunol*. 2005;42(5):651-655.

1456

3. Bostrom J, Haber L, Koenig P, Kelley RF, Fuh G. High affinity antigen recognition of the dual specific variants of herceptin is entropy-driven in spite of structural plasticity. *PLoS One*. 2011;6(4):e17887.
4. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. *Science*. 2003;301(5638):1374-1377.
5. Van Regenmortel MH. Specificity, polyspecificity, and heterospecificity of antibody-antigen recognition. *J Mol Recognit*. 2014;27(11):627-639.
6. Dimitrov JD, Planchais C, Roumenina LT, Vassilev TL, Kaveri SV, Lacroix-Desmazes S. Antibody polyreactivity in health and disease:
statu variabilis. *J Immunol*. 2013;191(3):993-999.
7. Jain D, Salunke DM. Antibody specificity and promiscuity. *Biochem J*. 2019;476(3):433-447.
8. Cohn M. A new concept of immune specificity emerges from a con- sideration of the self-nonself discrimination. *Cell Immunol*. 1997;181 (2):103-108.
9. Parnes O. From interception to incorporation: degeneracy and pro- miscuous recognition as precursors of a paradigm shift in immunol- ogy. *Mol Immunol*. 2004;40(14-15):985-991.
10. Ko S, Jung ST. Engineering antibodies for dual specificity and enhanced potency. *Biotechnol Bioprocess Eng*. 2015;20(2):201-210.
11. Chames P, Baty D. Bispecific antibodies for cancer therapy: the light at the end of the tunnel? *MAbs*. 2009;1(6):539-547.
12. Kaplon H, Reichert JM. Antibodies to watch in 2019. *MAbs*. 2019;11
(2):219-238.
13. Sedykh SE, Prinz VV, Buneva VN, Nevinsky GA. Bispecific antibodies:
design, therapy, perspectives. *Drug Des Dev Ther*. 2018;12:195-208.
14. Eigenbrot C, Fuh G. Two-in-one antibodies with dual action Fabs. *Curr Opin Chem Biol*. 2013;17(3):400-405.
15. Jiang G, Lee CW, Wong PY, Gazzano-Santoro H. Evaluation of semi- homogeneous assay formats for dual-specificity antibodies. *J Immunol Methods*. 2013;387(1-2):51-56.
16. Koenig P, Lee CV, Sanowar S, et al. Deep sequencing-guided design of a high affinity dual specificity antibody to target two angiogenic factors in neovascular age-related macular degeneration. *J Biol Chem*. 2015;290(36):21773-21786.
17. Koenig P, Sanowar S, Lee CV, Fuh G. Tuning the specificity of a two- in-one fab against three angiogenic antigens by fully utilizing the information of deep mutational scanning. *mAbs*. 2017;9(6):959-967.
18. Lee HY, Schaefer G, Lesaca I, Lee CV, Wong PY, Jiang G. "Two-in- one" approach for bioassay selection for dual specificity antibodies. *J Immunol Methods*. 2017;448:74-79.
19. Schaefer G, Haber L, Crocker LM, et al. A two-in-one antibody against HER3 and EGFR has superior inhibitory activity compared with monospecific antibodies. *Cancer Cell*. 2011;20(4):472-486.
20. Bostrom J, Yu SF, Kan D, et al. Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site. *Science*. 2009;323(5921):1610-1614.
21. Jain T, Sun T, Durand S, et al. Biophysical properties of the clinical- stage antibody landscape. *Proc Natl Acad Sci U S A*. 2017;114(5): 944-949.
22. Starr CG, Tessier PM. Selecting and engineering monoclonal anti- bodies with drug-like specificity. *Curr Opin Biotechnol*. 2019;60: 119-127.
23. Kaleli NE, Karadag M, Kalyoncu S. Phage display derived therapeutic antibodies have enriched aliphatic content: insights for developability issues. *Proteins*. 2019;87(7):607-618.
24. Chuang GY, Zhang B, McKee K, et al. Eliminating antibody poly- reactivity through addition of N-linked glycosylation. *Protein Sci*. 2015;24(6):1019-1030.
25. Igawa T, Tsunoda H, Tachibana T, et al. Reduced elimination of IgG antibodies by engineering the variable region. *Protein Eng Des Sel*.
2010;23(5):385-392.
26. Kelly RL, le D, Zhao J, Wittrup KD. Reduction of nonspecificity motifs in synthetic antibody libraries. *J Mol Biol*. 2018;430(1):119-130.
27. Rabia LA, Zhang Y, Ludwig SD, Julian MC, Tessier PM. Net charge of antibody complementarity-determining regions is a key predictor of specificity. *Protein Eng Des Sel*. 2018;31(11):409-418.
28. Dondelinger M, Filée P, Sauvage E, et al. Understanding the signifi- cance and implications of antibody numbering and antigen-binding surface/residue definition. *Front Immunol*. 2018;9:2278.
29. Rabia LA, Desai AA, Jhajj HS, Tessier PM. Understanding and over- coming trade-offs between antibody affinity, specificity, stability and solubility. *Biochem Eng J*. 2018;137:365-374.
30. Safdari Y, Farajnia S, Asgharzadeh M, Khalili M. Antibody humaniza- tion methods - a review and update. *Biotechnol Genet Eng Rev*. 2013; 29:175-186.
31. Wu H, Nie Y, Huse WD, Watkins JD. Humanization of a murine monoclonal antibody by simultaneous optimization of framework and CDR residues. *J Mol Biol*. 1999;294(1):151-162.
32. Haidar JN, Yuan QA, Zeng L, et al. A universal combinatorial design of antibody framework to graft distinct CDR sequences: a bioinformatics approach. *Proteins*. 2012;80(3):896-912.

33. Foote J, Winter G. Antibody framework residues affecting the confor- mation of the hypervariable loops. *J Mol Biol*. 1992;224(2):487-499.
34. Riechmann L, Clark M, Waldmann H, Winter G. Reshaping human antibodies for therapy. *Nature*. 1988;332(6162):323-327.
35. Almagro JC, Fransson J. Humanization of antibodies. *Front Biosci*. 2008;13:1619-1633.
36. Saldanha JW. Humanization strategies. In: Dübel SRJM, ed. *Handbook of Therapeutic Antibodies*. Weinheim, Germany: Wiley Blackwell; 2014.
37. Weitzner BD, Jeliazkov JR, Lyskov S, et al. Modeling and docking of antibody structures with Rosetta. *Nat Protoc*. 2017;12(2):401-416.
38. Lyskov S, Chou FC, Conchúir SÓ, et al. Serverification of molecular modeling applications: the Rosetta online server that includes every- one (ROSIE). *PLoS One*. 2013;8(5):e63906.
39. Kastritis PL, Bonvin AM. Are scoring functions in protein-protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark. *J Proteome Res*. 2010;9(5):2216-2225.
40. van Zundert GCP, Rodrigues JPGLM, Trellet M, et al. The HAD- DOCK2.2 web server: user-friendly integrative modeling of biomolec- ular complexes. *J Mol Biol*. 2016;428(4):720-725.
41. Spiliotopoulos D, Kastritis PL, Melquiond ASJ, et al. dMM-PBSA: a new HADDOCK scoring function for protein-peptide docking. *Front Mol Biosci*. 2016;3:46.
42. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372(3):774-797.
43. Madeira F, Park Y, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019;47(W1):W636- W641.
44. Shepherd P, Dean CJ. *Monoclonal Antibodies: A Practical Approach*. New York, NY: Oxford University Press; 2000.
45. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PROD- IGY: a web server for predicting the binding affinity of protein- protein complexes. *Bioinformatics*. 2016;32(23):3676-3678.
46. Vangone A, Bonvin AM. Contacts-based prediction of binding affinity in protein-protein complexes. *Elife*. 2015;4:e07454.
47. Pires DE, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics*. 2014;30(3):335-342.
48. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. *Nucleic Acids Res*. 2017;45(W1):W229-W235.
49. Pires DE, Ascher DB. mCSM-AB: a web server for predicting antibody-antigen affinity changes upon mutation with graph-based signatures. *Nucleic Acids Res*. 2016;44(W1):W469-W473.
50. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 2005;33:W306-W310.
51. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins*. 2006; 62(4):1125-1132.
52. Teplyakov A, Obmolova G, Malia TJ, et al. Structural insights into humanization of anti-tissue factor antibody 10H10. *MAbs*. 2018;10 (2):269-277.
53. Bond CJ, Wiesmann C, Marsters JC Jr, Sidhu SS. A structure-based database of antibody variable domain diversity. *J Mol Biol*. 2005;348 (3):699-709.
54. Fanning SW, Horn JR. An anti-hapten camelid antibody reveals a cryptic binding site with significant energetic contributions from a nonhypervariable loop. *Protein Sci*. 2011;20(7):1196-1207.
55. Henry KA, Hussack G, Kumaran J, et al. Role of the non-hypervariable FR3 D-E loop in single-domain antibody recognition of haptens and carbohydrates. *J Mol Recognit*. 2019;32(11):e2805.
56. Adib-Conquy M, Gilbert M, Avrameas S. Effect of amino acid substi- tutions in the heavy chain CDR3 of an autoantibody on its reactivity. *Int Immunol*. 1998;10(3):341-346.
57. Swindells MB, Porter CT, Couch M, et al. abYsis: integrated antibody sequence and structure-management, analysis, and prediction. *J Mol Biol*. 2017;429(3):356-364.
58. Birtalan S, Fisher RD, Sidhu SS. The functional capacity of the natural amino acids for molecular recognition. *Mol Biosyst*. 2010;6(7):1186-1194.
59. Dalkas GA, Teheux F, Kwasigroch JM, Rooman M. Cation-pi, aminopi, pi-pi, and H-bond interactions stabilize antigen-antibody inter- faces. *Proteins*. 2014;82(9):1734-1746.
60. Schreiber G, Keating AE. Protein binding specificity versus promiscu- ity. *Curr Opin Struct Biol*. 2011;21(1):50-61.
61. Peng HP, Lee KH, Jian JW, Yang AS. Origins of specificity and affinity in antibody-protein interactions. *Proc Natl Acad Sci U S A*. 2014;111 (26):E2656-E2665.
62. Makabe K, Nakanishi T, Tsumoto K, et al. Thermodynamic conse- quences of mutations in Vernier zone residues of a humanized anti- human epidermal growth factor receptor murine antibody, 528. *J Biol Chem*. 2008;283(2):1156-1166.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Arslan M, Karadag D, Kalyoncu S. Conformational changes in a Vernier zone region: Implications for antibody dual specificity. *Proteins*. 2020;88:1447-1457.
https://doi.org/10.1002/prot.25964

scientific reports

Check for updates

OPEN

Effect of non-repetitive linker on in vitro and in vivo properties of an anti-VEGF scFv

Merve Arslan[1,2], Murat Karadag[1,2], Ebru Onal[1,3], Emine Gelinci[1], Gulcin Cakan-Akdogan[1,4] & Sibel Kalyoncu[1]✉

Single chain antibody fragments (scFvs) are favored in diagnostic and therapeutic fields thanks to their small size and the availability of various engineering approaches. Linker between variable heavy (V_H) and light (V_L) chains of scFv covalently links these domains and it can affect scFv's bio-physical/chemical properties and in vivo activity. Thus, scFv linker design is important for a successful scFv construction, and flexible linkers are preferred for a proper pairing of V_H–V_L. The flexibility of the linker is determined by length and sequence content and glycine-serine (GS) linkers are commonly preferred for scFvs based on their highly flexible profiles. Despite the advantage of this provided flexibility, GS linkers carry repeated sequences which can cause problems for PCR-based engineering approaches and immunogenicity. Here, two different linkers, a repetitive GS linker and an alternative non-repetitive linker with similar flexibility but lower immunogenicity are employed to generate anti-Vascular Endothelial Growth Factor scFvs derived from bevacizumab. Our findings highlight a better in vitro profile of the non-repetitive linker such as a higher monomer ratio, higher thermal stability while there was no significant difference in in vivo efficacy in a zebrafish embryonic angiogenesis model. This is the first study to compare in vivo efficacy of scFvs with different linkers in a zebrafish model.

Monoclonal antibodies (mAbs) and antibody fragments are used in a variety of therapeutic and diagnostic applications[1]. Due to its small size and availability of various protein engineering techniques, single chain variable fragment (scFv) is one of the most utilized antibody fragments. scFv is composed of variable domains of heavy (V_H) and light (V_L) chains of mAbs that are covalently linked together by a "exible peptide linker. Linker nature is key to forming a proper V_H–V_L antigen-binding interface a#ecting scFv function. scFv linker is critical due to its e#ects on both its in vitro and in vivo properties[2,3]. $erefore, peptide linker design is key for a successful scFv construction[4,5].

$e length and sequence content of the linker are two features that can a#ect expression level, folding, oligomeric state, a%nity/speci&city, stability, and in vivo activity of scFvs[5–8]. Natural and synthetic linkers are being studied for fusion proteins and they are broadly divided into three groups: (i) "exible, (ii) rigid, and (iii) cleavable[7,9]. Glycine-serine (GS) repeat is the most common linker sequence in scFv design mostly due to their "exible nature. GS linkers are utilized in some of the very &rst scFv fragments[10]. Di#erent lengths and combinations of GS linkers are tested for scFv fragments, the most common ones are $(G_3S)_n$ and $(G_4S)_n$ motifs[3,11–13]. $e length of the linker can be optimized from 5 to 35 amino acids to develop improved scFvs for various applications[5,14]. If the scFv linker length is longer than 12 residues, covalently linked V_H and V_L form a functional scFv and they are supposed to be highly monomeric. scFvs with shorter linkers (< 12 a.a.) tend to form multimers by combining with other scFv molecule(s)[15,16].

Although GS linkers provide the "exibility which is desired for proper scFv folding and structure, they are repetitive which can cause problems related to PCR-based engineering strategies[17] and immunogenicity[18]. $erefore, alternative non-repetitive linkers with comparable "exibility can be employed for improved in vivo properties. In this study, we compared in vitro and in vivo properties of repetitive and non-repetitive linker sequences utilized in an anti-Vascular Endothelial Growth Factor (VEGF) scFv. $e role of VEGF is critical as it is a key driver of the sprouting angiogenesis in tumor growth[19]. $is has led to the development of anti-VEGF therapeutic approaches and many therapeutic drugs against VEGF are being used[20,21]. $is is the &rst study to

[1]Izmir Biomedicine and Genome Center, Izmir, Turkey. [2]Izmir International Biomedicine and Genome Institute, Dokuz Eylul University, Izmir, Turkey. [3]Institute of Health Sciences, Dokuz Eylul University, Izmir, Turkey. [4]Department of Medical Biology, Faculty of Medicine, Dokuz Eylul University, Izmir, Turkey. ✉email: sibel.kalyoncu@ibg.edu.tr

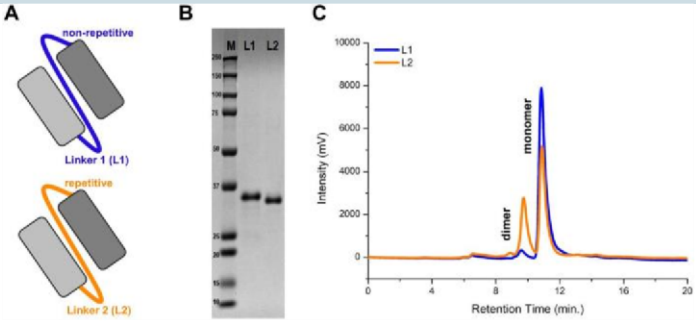


Figure 1. (**A**) Schematic representation of designed L1 and L2 scFvs. (**B**) SDS-PAGE analysis of puri$ed L1 and L2. Full image of the SDS-PAGE gel is provided in Supplementary Figure 1. (**C**) SEC chromatograms of L1 (blue) and L2 (orange) from SE-HPLC analysis.

	Flexibility score	Immunogenicity score	MW (kDa)	Monomer/dimer ratio (monomer %)	% Insoluble aggregates*
L1	0.47	−1.09	32.8	95.7 ± 0.7	88.5 ± 2.8
L2	0.53	−0.06	31.9	66.5 ± 3.2	90.2 ± 4.3

Table 1. Properties of L1 and L2 variants. *Percent insoluble aggregation was calculated by subtracting soluble protein concentration from total concentration a˜er thermal (60 °C) and mechanical (220 rpm) stress for 4 h. &e average of 3 di#erent samples was used.

compare in vivo e''cacy of scFvs with di#erent linkers on a zebra$sh model. Although non-repetitive linker showed better in vitro properties, there was no signi$cant di#erence in in vivo e''cacy. More importantly, both of our designed scFvs which are derived from the bevacizumab sequence showed better in vivo e#ects than bevacizumab itself.

Results

Two di#erent linker sequences for the same anti-VEGF scFv sequence were used. Both linkers have long lengths for proper structural positioning of linked V_H and V_L to form monomeric scFv. While one scFv has a non-repetitive linker (L1, SPNSASHSGSAPQTSSAPGSQ), the other has a repetitive linker (L2, $(G_3S)_4$) (Fig. 1A). First, we performed in silico analysis to determine the %exibility of the linkers and it showed that they have very similar %exibilities (Table 1). Immunogenicity propensities were also investigated by a speci$c IEDB tool[22]. A higher score implies a bigger probability of eliciting an immune response. L2, scFv with a repetitive linker, showed a signi$cantly higher immunogenicity score than L1. Because immunogenicity is undesired in therapeutic interventions, this particular non-repetitive linker might be advantageous compared to the most common GS linker.

L1 and L2 were expressed in *E.coli* and puri$ed from the supernatant by Protein L chromatography with high purity. &e molecular weight di#erence of ~ 1 kDa due to the linker sequence di#erence can be distinguished in SDS-PAGE (Fig. 1B). &e e#ect of the linkers on the oligomeric state of the scFvs was investigated using size-exclusion chromatography (SEC). Retention times of dimeric and monomeric forms were identi$ed using molecular weight protein standards. Percentages of dimeric and monomeric forms in solution were determined using areas under the peaks. According to the results, there is a signi$cant di#erence in monomer/dimer ratios (Table 1, Fig. 1C). L1 has > 95% monomer while L2 has 66.5% monomer. A higher monomer ratio is more desired because it is known to be thermodynamically more stable[23].

Aggregation propensities of L1 and L2 were examined under stress conditions. Aggregation of the proteins was induced by both mechanical (shaking) and thermal (heating) stress and soluble fractions of the proteins a˜er precipitation of aggregate forms were measured. &ere was no signi$cant di#erence in aggregation pro$les of L1 and L2 (Table 1). &ermal stabilities of L1 and L2 were determined by thermal denaturation assay. A %uorescent dye that binds to hydrophobic regions was used to monitor protein unfolding under thermal stress. Melting temperature (T_m) at which the half percentage of protein is unfolded was calculated. L1 was more stable compared to L2 indicating that non-repetitive linker provided more thermal stability to the scFv (Fig. 2).

Binding kinetics of L1, L2, and bevacizumab (IgG) were analyzed based on their bindings to their ligand, VEGF. Surface Plasmon Resonance (SPR) was used and corresponding association (k_{on}), dissociation ($k_{o\#}$) constants, and binding a''nity (K_D) were obtained (Fig. 3). A comparison of the kinetic parameters is listed in Fig. 3A. Binding a''nities were determined as 0.38 nM, 2.51 nM and 0.83 nM for Bevacizumab, L1 and L2,

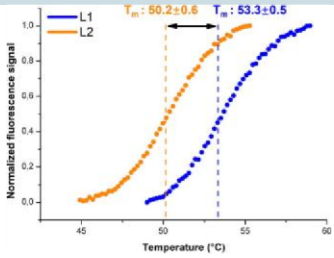


Figure 2. ▪▪ermal melting temperatures of L1 and L2. Transition mid-points (T_m values) from ▪uorescent thermal melt assays were calculated by Hill equation #t. ▪▪e assay was repeated 3 times.

A

	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
Bevacizumab	2.6×10^5	1.0×10^{-4}	0.38 ± 0.02
L1	3.9×10^4	9.9×10^{-5}	2.51 ± 0.01
L2	5.7×10^4	4.7×10^{-5}	0.83 ± 0.02

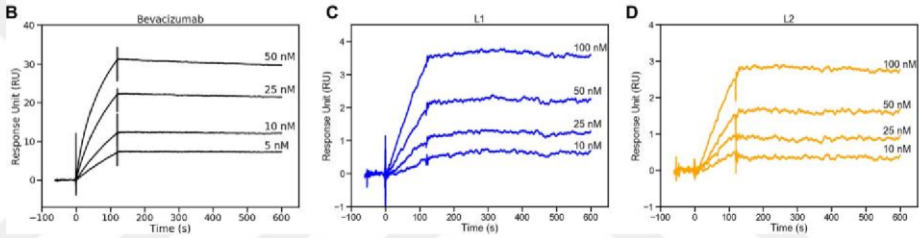


Figure 3. Binding kinetics of Bevacizumab, L1 and L2 fragments to their antigen, VEGF. (**A**) Obtained kinetic parameters. Sensogram overlays for (**B**) bevacizumab (**C**) L1 and (**D**) L2. Antibody concentrations are expressed on corresponding curves. Analyses were repeated 2 times.

respectively. Both scFvs have close a▪nities to Bevacizumab. L2 showed approximately threefold better binding to VEGF compared to L1. ▪▪ signi#cant di$erence in oligomeric states of scFvs caused by di$erent linkers might be e$ective on this di$erence in binding characteristics.

In vivo anti-angiogenesis activities of L1 and L2 were tested with the zebra#sh subintestinal vessel (SIV) assay, using *▪i1:EGFP* transgenic zebra#sh line which drives GFP expression in endothelial cells[24]. L1, L2, Bevacizumab, or PBS was injected into the yolk of embryos at 48–52 h post fertilization (2dpf) at the initiation phase of SIV formation[25]. ▪▪e SIV development is driven by VEGF in zebra#sh the anti-VEGF agents and antibodies were shown to inhibit SIV development in zebra#sh[26]. Antibodies L1 (55 &M), L2 (55 &M) and bevacizumab (27.5 &M) were injected into the yolk of 2 dpf (48–52 hpf) zebra#sh embryo, the e$ect was quanti#ed the next day by analyzing the SIV area. Half concentration for bevacizumab was used because bevacizumab is divalent (1:2 antibody: VEGF binding) and scFvs are monovalent (1:1).

SIVs of embryos treated with 1X PBS were ordered and intact as expected, whereas L1 and L2 inhibited SIV development in the majority of embryos (Fig. 4). To quantify the inhibition, the SIV area in each embryo was measured and the relative mean area with respect to the control group was depicted as percentages (Fig. 4F). When compared to 1X PBS control, L1 and L2 induced a decrease in SIV areas by 26% and 34.4%, respectively (Fig. 4F). However, bevacizumab did not show a signi#cant e$ect at 27.5 &M. When a higher dose (55 &M) of bevacizumab was injected, a 35.6% decrease in SIV area was observed, similar to that of L2 (data not shown).

Discussion

Linker contents of scFv are known to a$ect their in vitro and in vivo properties, so linker design is critical for the developability of an scFv. $(G_nS)_n$ linkers are the most common linkers due to their shown ▪exibility. In a comparative study performed by Vihinen, *et. al.*, amino acids with bulky side chains such as lysine, and aspartate also have higher ▪exibility[27]. ▪▪e repetitive sequence of the GS linker can cause issues during the introduction of random

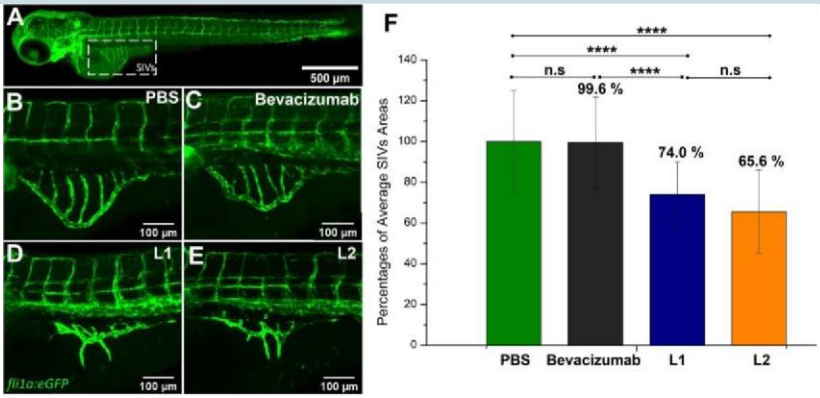


Figure 4. In vivo angiogenesis inhibition by L1, L2 and bevacizumab. (**A**) Lateral view of subintestinal vessels (SIVs) in *fli1:EGFP* transgenic zebra"sh larvae at 3 dpf. (**B**) PBS (negative control) (**C**) 27.5 *M bevacizumab (**D**) 55 *M L1, (**E**) 55 *M L2 were injected into yolk of 2 dpf *fli1:EGFP* transgenic zebra"sh embryos. At 3 dpf, zebra"sh SIVs were imaged by confocal microscopy. (**F**) Percentages of average SIV areas were quanti"ed. $t outcomes are expressed as AVG ± SD. $n_{PBS} = 25$, $n_{bevacizumab} = 24$, $n_{L1} = 24$, $n_{L2} = 23$. Statistical analysis was performed using one-tail t-test. Statistical results: n.s. $p > 0.05$, ****$p < 0.0001$.

mutations via mutagenic DNA shu!ing or during PCR-based assembly of ampli"ed variable domains, which may result in undesired length variants due to improper annealing of homologous sequences of these repetitive sequences[28,29]. $ere is also a high possibility of having a higher immunogenic e%ect of repetitive linkers due to their O-glycosylation patterns[30]. Plückthun, et al. addressed these issues by designing and selecting non-repetitive linker sequences without disrupting desired properties such as proper folding, solubility, and binding by using the selectively infective phage technology[17]. $ree non-repetitive linkers were chosen and further characterized in that work and shown to be equivalent to the original scFv fragment with the $(G_4S)_3$ linker.

Although linker design and engineering are very important in scFv construction, alternative approaches to GS linkers are very limited in the literature. Klement, et al. screened natural and arti"cial linkers for their cytotoxic antibody fragment[31]. $ey showed that structural stability and functionality were signi"cantly a%ected by linker content and a natural linker composed of IgG3 upper hinge region had the highest functionality and stability. In one study, three di%erent GS-bearing linkers with neutral, positive, and negative charges were employed and no signi"cant di%erence was observed for their biodistribution patterns in tumor-bearing mice[32]. In another study, GS and only glycine-containing linkers varying from 1 to 25 amino acids were screened and GS linkers showed better biophysical characteristics and pharmacokinetic properties[14].

Here, we employed two distinct types of linkers to design an anti-VEGF scFv antibody derived from bevacizumab[33]. One is the most preferred &exible and repetitive GS linker ($(G_3S)_4$, L2), and the other is a non-repetitive linker of "SPNSASHSGSAPQTSSAPGSQ" (L1)[17]. While L2 is a simple linker with repeated glycine and serine residues, it might lead to reduced stability and/or higher immunogenicity which is undesired in therapeutic antibody development. First, developed scFvs were compared to each other in terms of their immunogenicity and &exibility scores using in silico tools. Although their &exibilities were similar, the non-repetitive linker was predicted to be less immunogenic. Both L1 and L2 have linkers with su➢cient length and &exibility which should allow both scFv antibodies to be predominantly monomeric structure[34]. Shorter linkers (less than 15 amino acids) enhance oligomerization due to pairing constraints of covalently linked V_H and V_L domains[35]. Higher-order oligomers of scFvs tend to have decreased overall stability and higher aggregation propensities[14,36]. $us, both linkers which are longer than 15 amino acids should allow the domains to pair in the proper orientation for dominantly monomeric forms. First, we analyzed oligomeric states of L1 and L2 by SE-HPLC. As expected, both scFvs were dominantly monomeric. However, scFv with the non-repetitive linker had a signi"cantly higher monomeric form which is desired (95.7% and 66.5% for L1 and L2, respectively). By using in vitro assays, we also compared the thermal stability, aggregation resistance, and a➢nity characteristics of these two scFvs. $ey did not show a signi"cant di%erence in terms of their aggregation tendencies a(er thermal and mechanical stress. While L1 had signi"cantly higher thermal stability, it had ~ 3 × less a➢nity to VEGF. It is important to note that VEGF a➢nities of L1 and L2 were close to bevacizumab which are all within the therapeutic range. (Fig. 3). $is shows that the binding a➢nity of engineered scFvs is usually preserved upon exclusion of constant regions[37]. $ese in vitro results show that the structural stability of the scFv with the non-repetitive linker is better with a slight decrease in binding functionality.

In vivo e➢cacy of these scFv antibodies were compared with bevacizumab using the zebra"sh (*Danio rerio*) SIV assay. According to our knowledge, this is the "rst study investigating the linker e%ect on scFv antibodies by using a zebra"sh model. Zebra"sh SIV development is induced by VEGF, and previous studies showed

that human anti-VEGF antibodies can inhibit this process[38]. Due to the transparency of the zebra!sh and, the robustness of the embryo and larvae for micromanipulation, zebra!sh is an advantageous model for in vivo testing of anti-VEGF antibodies[39,40]. In this study, we tested the in vivo e#cacies of L1 and L2 scFvs and compared them to that of bevacizumab. $here was no systemic deformation observed a%er microinjections. Both L1 and L2 scFvs were found to be more e&ective than bevacizumab when applied at the same stoichiometric ratio with respect to VEGF binding. Although bevacizumab is divalent and can bind to VEGF at a 1:2 ratio, and scFvs are monovalent and bind to VEGF at a 1:1 ratio, the half concentration of bevacizumab was not as e&ective as scFvs. When 55 ˚M of L1 or L2 were injected, SIV development was inhibited which resulted in an irregular structure and a reduction in the area of SIVs. According to statistical comparisons with the control group (1X PBS), 26% and 34.4% reductions were observed for L1 and L2, respectively. However, bevacizumab as positive control did not show a signi!cant e&ect at 27.5 ˚M. When the bevacizumab concentration was increased 2 × to 55 ˚M, it caused a 35% reduction in the SIV area, a comparable activity to our scFvs. Considering that scFv is a fragment of bevacizumab, it is expected to be as e&ective as or more e&ective than bevacizumab. Although bevacizumab had similar in vivo e#cacy when used as 2 × more stoichiometric ratio concentration, it can be concluded that bevacizumab was less e&ective than designed scFvs. It is known that smaller antibody fragments have better tissue penetration[41–43], so better e#cacy of our designed scFvs in the zebra!sh model can be explained based on their !vefold smaller sizes (~ 30 kDa) compared to parent bevacizumab (~ 150 kDa). Also, the biodistribution coe#cient (estimation of tissue distribution based on plasma concentration) of smaller size antibody fragments was found to be higher which would increase drug e#cacy in targeted tissue[44]. E&ective scFvs can be derived from commonly used IgGs with several protein engineering techniques[10,45–49]. Overall, this study shows that scFvs can be potentially used for diagnostic and therapeutic approaches and linker design is important for some of their key developability characteristics.

Methods

Genes and protein expression. Anti-VEGF scFv sequence was derived from bevacizumab. Bevacizumab sequence was extracted from the DrugBank database with the accession number DB00112[50]. Antibody residues were numbered according to the Kabat numbering system. Variable domain residues were determined via analysis of PDB structure of bevacizumab, 1BJ1[51]. C-terminal of the variable domains sequences are determined a%er the last β-sheet and the !rst 2 residues of the linker region for the heavy and light chain. $ese determined heavy ($V_H$) and light ($V_L$) chain sequences were linked via either a non-repetitive or repetitive linker. scFv with non-repetitive linker (L1) and repetitive linker (L2) has linker sequences of "SPNSASHSGSAPQTSSAPGSQ" and $(G_3S)_4$, respectively. L1 sequence was selected from the study of Hennecke, et al. with one amino acid di&erence from the reference linker (glutamine instead of asparagine at position 13) which gave lower immunogenicity and higher *exibility score[17]. scFv variants with a leader sequence (PelB), FLAG-tag and penta histidine-tag were transformed into *E.coli strain BL21 (DE3) pLysS* ($ermo Fisher) with pET17-b (GenScript) expression plasmid. Transformant cells were grown on LB-agar plate containing 100 ˚g/mL ampicillin and 25 ˚g/mL chloramphenicol. Single colonies were inoculated in LB broth containing 100 ˚g/mL ampicillin and 25 ˚g/mL chloramphenicol and grown overnight at 225 rpm, 37 °C as inoculum. $ese cells were inoculated into 300 mL autoinduction-media and incubated at 18 °C, 250 rpm for 48 h[52].

Protein purification. Cultures were centrifuged at 6500xg at 4 °C (Avanti, Beckman Coulter). Protein-containing supernatant was incubated with His-Pur Ni-NTA resin ($ermo Fisher) for 2 h at 4 °C, mixing gently. $e mixture was loaded onto a 10 ml vacuum column ($ermo Fisher) and puri!ed according to the manual's protocol. 1X phosphate-bu&ered saline (PBS) with 25 mM imidazole, pH 7.4 and PBS with 500 mM imidazole, pH 7.4 were used as a wash and elution bu&ers, respectively. Puri!ed proteins were bu&er-exchanged[53] into PBS (pH 7.4) through membrane !ltration (Amicon® Ultra-4 Centrifugal Filter Units, MWCO 10 kDa, Merck). Protein samples were then loaded onto the HiTrap™ Protein L column (GE Healthcare) to achieve better purity (> 95%). Protein purities were con!rmed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis. Precision Plus Protein™ Dual Color standard was used as a marker (Bio-Rad). Protein concentration was determined by NanoDrop 2000 (absorbance at 280 nm).

Computational tools. *Flexibility analysis.* $e average *exibility of non-repetitive and repetitive linker sequences are calculated according to the average *exibility index of amino acids[54,55]. Flexibility value calculations for L1 and L2 were done based on their amino acid composition and length.

Immunogenicity predictions. $e risk of immunogenicity of L1 and L2 was predicted using an online tool from IEDB T Cell Epitopes Immunogenicity Prediction (http://tools.immuneepitope.org/main/)[22].

Size-exclusion chromatography (SE-HPLC). A Shimadzu Prominence HPLC System with UV–VIS detector (Kyoto, Japan), and an analytical TSK-gel SuperSW3000 column (4.6 mm ID x 30 cm, 4 mm) (Tosoh Bioscience, Tosoh, USA) were used to perform SEC. Running bu&er was 0.1 M phosphate bu&er, and 0.1 M potassium sulfate (pH 6.7). scFv samples were prepared at 0.1 mg/mL and injected into the column at 25 °C with a *ow rate of 0.3 mL/min over 20 min. $e absorbance values were monitored at 280 nm. $yroglobulin ($M_r$ = 669,000), IgG (M_r = 150,000), BSA (M_r = 66,400), myoglobin (M_r = 17,000), and uracil (M_r = 112) were used as molecular weight standards to verify the retention times of the scFv samples.

Aggregation analysis. Protein aliquots (0.5 mg/mL, 25 "L) were incubated at 60 °C, 220 rpm in a heat block to provide both thermal and mechanical stress to seed aggregation. At di#erent time intervals (0–420 min), protein aliquots were taken and centrifuged at 17,000xg at 4 °C for 10 min to pellet aggregated part. Protein concentrations of soluble fractions were measured using NanoDrop 2000 (absorbance at 280 nm).

Thermal denaturation assay. $ermal unfolding pro%les of puri%ed scFv proteins were determined with SYPRO Orange dye (Sigma, S5692) via ABI 7500 Fast RT-PCR (25–99 °C with 0.05% ramp rate). Optimum concentrations of dye and protein were 2 × and 2 "M, respectively. Transition mid-points (T_m values) from the thermogram data were calculated using the Hill1 equation %t using the Origin 8.5 so"ware.

Surface plasmon resonance (SPR). A(nity measurements were performed using a Biacore T200 instrument (Biacore Inc., Piscataway, NJ). All experiments were performed in HBS-EP bu#er, pH 7.4. 1000 nM VEGF protein was immobilized onto a CM4 chip (Cytiva) at a)ow rate of 10 µl/min for ~1 min. A series of solutions ranging from 10 to 100 nM scFvs and 10 to 50 nM bevacizumab were subsequently injected at a)ow rate of 30 µl/min onto the VEGF-coated surface. Data were corrected by double-referencing against a control)ow cell containing no VEGF and against the)ow cell with bu#er injection. Sensogram curves were analyzed using the BiaEval 3.0 manufacturer's so"ware. $e dissociation constant (K_D), association rate constant (k_{on}) and dissociation rate constant ($k_{o#}$) values were calculated by %tting the kinetic association and dissociation curves to a 1:1 binding model.

Zebrafish experiments. Zebra%sh used in this study were provided by the Zebra%sh Facility in Izmir Biomedicine and Genome Center. All animal procedures were approved by the IBG Local Ethics Committe for Animal Experimentation (IBG-HADYEK) with protocol no 2020-013.

Adult zebra%sh were maintained under standard conditions at 28 °C on a 14/10-h light/dark cycle, at the Izmir Biomedicine and Genome Center Zebra%sh Facility. Transgenic *fli1:EGFP* line[24] was crossed to wild type AB to obtain the embryos, which were collected within 30 min of fertilization, incubated in E3 embryo medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$,0.33 mM $MgCl_2$, %1 methylene blue) at 28 °C in a dark incubator.

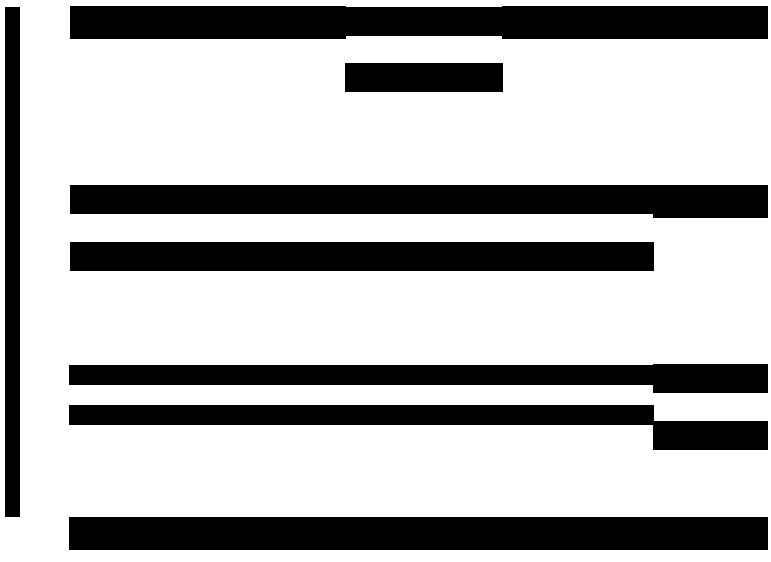
Microinjection and #xation. 55 "M L1, L2 and 55 or 27.5 "M bevacizumab (in PBS and 0.5% Phenol red) were kept on ice until microinjection. $e estimated injection volume was determined by the use of a stage micrometer and 50 nL of L1, L2, and bevacizumab were microinjected into the yolk of each anesthetized zebra%sh embryo (0.02% tricaine) at 2 days post fertilization (dpf) (n > 20 embryos per condition). Microinjected larvae were kept at 28 °C and at 1-day post-injection (dpi), development of subintestinal vessels (SIVs) were examined. Larvae were %xed with fresh 4% formaldehyde, washed with PBS, melanocytes were decolorized with pigment discoloration solution (5% of KOH, 30% of H_2O_2, PBS (1:1:8)). Samples were kept at 4 °C until imaging.

Imaging and quanti#cation. Larvae were mounted laterally in 1% low melting agarose in imaging dish (D35-14–1.5 N Cell Vis, USA), imaged with Zeiss LSM880 confocal microscope. SIV area was measured with ImageJ polygon tool, statistical analysis was performed with Excel using the one-tail t-test with a con%dence interval of 95%. Statistical signi%cances were represented with stars of p-values: non-signi%cant (ns) > 0.05, *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001. $e outcomes are expressed as average (AVG) ± standard error of the mean (SEM).

Received: 24 November 2021; Accepted: 3 March 2022

Published online: 31 March 2022

References



[redacted]

Acknowledgements

Authors would like to acknowledge Izmir Biomedicine and Genome Center (IBG Startup Grant), YÖK (Turkish Council of Higher Education) 100/2000 Fellowship Program, TUBITAK-BIDEB 2211-A PhD Achievement Grant, and TUBITAK grant no 119Z161 for funding this research. Authors thank IBG Optical Imaging Core Facility manager Melek Ucuncu for support and Zebra!sh Facility sta" Meryem Ozaydin for excellent !sh care. We thank Dilara Karadağ for her invaluable in in vitro analysis.

Author contributions

M.A., M.K., E.O., G.C.A. and S.K. contributed to conception and design of the study. M.A. and M.K. carried out the in vitro experiments. E.O. and E.G. carried out the in vivo experiments. M.A., M.K. and E.O. wrote the !rst dra% of the manuscript. M.A., G.C.A. and S.K. contributed to manuscript revision. M.A., M.K., E.O., E.G., G.C.A. and S.K. read and approved the submitted version. G.C.A. and S.K. supervised the project.

Competing interests

&e authors declare no competing interests.

Additional information

Supplementary Information &e online version contains supplementary material available at https://doi.org/10.1038/s41598-022-09324-4.

Correspondence and requests for materials should be addressed to S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional a˜liations.

BBA - Proteins and Proteomics 1871 (2023) 140915

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

journal homepage: www.elsevier.com/locate/bbapa p

Research Paper

Engineering of conserved residues near antibody heavy chain complementary determining region 3 (HCDR3) improves both affinity and stability

Merve Arslan [a,b,1], TugVçe Uluçay [a,1], Seyit Kale [a], Sibel Kalyoncu [a,*]

[a] Izmir Biomedicine and Genome Center, Balçova, 35340 Izmir, Turkey

[b] Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, Balçova, 35340 Izmir, Turkey

ARTICLEINFO

Keywords:
Antibody
Antibody engineering
Vernier zone
Affinity-stability trade-offs
Molecular dynamics

ABSTRACT

Affinity and stability are crucial parameters in antibody development and engineering approaches. Although improvement in both metrics is desirable, trade-offs are almost unavoidable. Heavy chain complementarity determining region 3 (HCDR3) is the best-known region for antibody affinity but its impact on stability is often neglected. Here, we present a mutagenesis study of conserved residues near HCDR3 to elicit the role of this region in the affinity-stability trade-off. These key residues are positioned around the conserved salt bridge between V_H-K94 and V_H-D101 which is crucial for HCDR3 integrity. We show that the additional salt bridge at the stem of HCDR3 (V_H-K94:V_H-D101:V_H-D102) has an extensive impact on this loop's conformation, therefore simultaneous improvement in both affinity and stability. We find that the disruption of t-t stacking near HCDR3 (V_H-Y100E:V_L-Y49) at the V_H-V_L interface cause an irrecoverable loss in stability even if it improves the affinity.

Molecular simulations of putative rescue mutants exhibit complex and often non-additive effects. We confirm that our experimental measurements agree with the molecular dynamic simulations providing detailed insights for the spatial orientation of HCDR3. V_H-V102 right next to HCDR3 salt bridge might be an ideal candidate to overcome affinity-stability trade-off.

1. Introduction

Antibodies are widely utilized as diagnostic and therapeutic tools thanks to their high affinity and specificity towards target antigens. Hundreds of therapeutic antibodies in different formats such as Fab, single-chain variable fragment (scFv), or nanobodies, continue being developed and tested in clinical trials [1,2]. Due to their small size, monovalent nature, and simpler folding paths [3], scFvs offer several beneficial attributes for *in silico* and *in vitro* developability applications [4–6]. Although scFvs might show some drawbacks such as lower sta- bility, and fast clearance, they are readily suitable for protein engi- neering strategies to overcome them [7,8]. The eventual success of the antibody depend strongly on its developability properties such as ac- tivity and stability. The activity of an antibody is related to its strong and specific binding to its cognate antigen, a metric that is known as *antigen affinity* [9]. A molecular-level understanding of the affinity is a chal- lenging task that might limit antibody development. Affinity improvement efforts can result in detrimental effects on the overall profile [10,11]. The primary bottleneck here is *stability* which is another important developability parameter often negatively correlated with the affinity [12–14]. For a successful antibody, it is critical to maintain a necessary level of stability while engineering the affinity. Thus, co- screening of stability and affinity is essential to find the optimal muta- tions [15,16].

Affinity modulation efforts have historically focused on the complementarity determining regions (CDRs), particularly heavy chain CDR3 (HCDR3) [17,18]. These regions are intrinsically *hypervariable* originating from somatic hypermutations in mature B cells which lead to CDR sequence diversity [19]. Several studies point out that non-CDR regions (known as framework residues) can also be important for af- finity because they impact the overall antibody structure [11,20,21]. Affinity gains by substitutions of non-CDR residues have been generally underestimated in directed evolution strategies. And there has been little systematic study to assess the consequences of this bias. In this

* Corresponding author at: Izmir Biomedicine and Genome Center, Balçova, 35340 Izmir, Turkey. *E-mail address:* sibel.kalyoncu@ibg.edu.tr (S. Kalyoncu).

[1] The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

https://doi.org/10.1016/j.bbapap.2023.140915
Received 16 December 2022; Received in revised form 11 April 2023; Accepted 11 April 2023
Available online 13 April 2023

context, one of the non-CDR regions is the Vernier zone [22]. The Vernier zone of an antibody is characterized by a set of critical framework residues underlying the CDRs [23,24]. These residues potentially affect the conformations of the CDR loops and their orientation with respect to the antigen epitope; thus, they are common targets of humanization efforts to regain affinity via back-mutations [22,25,26]. Roles of the Vernier zone region on other antibody properties such as specificity and stability remain to be elucidated [27].

Powerful experimental technologies such as directed evolution and rational designs have made *in vitro* evolution of antibodies more efficient and have taken the field of modern antibody engineering further [28]. While experimental approaches can generally provide reliable results, scalability can be cumbersome, labor-intensive, and prohibitively expensive [29]. In this context, *in silico* approaches can provide invaluable opportunities with their high-throughput potential and ready access to atomic-level details [30]. Molecular dynamics (MD) simula- tions investigate the biomolecular structures of antibodies at their nat- ural dynamics on timescales relevant to their physiological function. These efforts can provide critical supporting information, including (i) computed free energy differences and measured forces behind protein- protein binding [31,32], (ii) atomic-level dynamics, (iii) information about protein stability in different physiological and experimental conditions [33–35], and (iv) the nature of the interactions between an antibody and its cognate antigen [36]. MD methods combined with rational design approaches can be used to design better antibodies in a shorter time with improved accuracy, and at a reduced cost [37]. The number of successful examples of this approach published in literature is on the rise [38–43].

In this study, we generated rationally chosen mutations on an anti- Vascular Endothelial Growth Factor (VEGF) scFv to investigate their effects on affinity and stability. Salt bridges are one of the most critical non-covalent forces in protein structure and function [44]. The salt bridge between the two heavy-chain residues 94 and 101 (according to Kabat numbering) is a highly conserved structural motif that supports the robust shape of HCDR3 [45,46]. Modifying or altering this interac- tion is virtually always detrimental for both stability and affinity [47]. Motivated by this critical observation, we hypothesized that the vicinity of this salt bridge is a natural starting point to investigate the trade-off between affinity and stability. To this end, we generated an anti-VEGF scFv antibody [48] and designed mutations in both heavy chain (V_H) and light chain (V_L) sides of this salt bridge accompanying the Vernier zone residues around (Fig. 1A). Anti-VEGF antibodies are successful anti-cancer therapeutics [49] and thanks to its universal nature of the Vernier zone, the lessons learned in this study could be transferable to other antibodies in general. We performed both comparative molecular dynamics simulations and experimental characterizations to gain a molecular-level understanding of the factors that modulate this trade- off. We found that existing and *de novo* secondary/tertiary interactions around this HCDR3 salt bridge are a critical determinant of both antigen binding and the robustness of the V_H-V_L interface, thus playing a crucial and complex role in the co-evolution of affinity-stability. Our overall findings obtained from our experimental and MD studies show the importance of joint efforts to elucidate the molecular mechanisms of antibody design.

2. Materials and methods

2.1. Setup preparation and analysis of molecular dynamics simulations

Atomic coordinates of the anti-VEGF scFv antibody fragment were taken from the Protein Data Bank (PDB ID:1BJ1; chains H, L and W) [50]. Constant fragment groups and the linker between the variable groups were omitted. When scFv linker length is longer than 12 residues, covalently linked V_H and V_L form a functional and monomeric scFv [51–53] and it was previously shown that absence of linker does not affect calculated distributions of molecular dynamic simulations [54].

Amino acid distributions of each Kabat numbering positions were extracted from AbYsis database (www.abysis.org). Mutant constructs were prepared using the Wizard Mutagenesis tool of PyMOL Molecular Graphics System, version 2.1.1, Schro·dinger LLC. Each antibody-antigen complex is solvated in a cubic water box that is sufficiently large to provide a minimum buffer zone of 12 Å between biological material and the cubic system boundaries. Na_+ and Cl_- ions were placed randomly to neutralize the system electrostatically at a physiological salt concentration of 0.150 M. CHARMM36m force field [55,56] was chosen together with the four-site OPC water model [57] subject to periodic boundary conditions. A combination of conjugate gradient and steepest descent methods were applied for initial energy minimization. Later, the system was equilibrated in the NVT ensemble at 100 K for 1 ns, and at 310 K for 1 ns, both using a small integration time step of 1 fs. Pro- duction trajectories were collected in the NPT ensemble at 310 K and 1 atm atmospheric pressure using a 2 fs of integration time steps for a total of 500 ns. Atomic coordinates were saved every 100 ps.

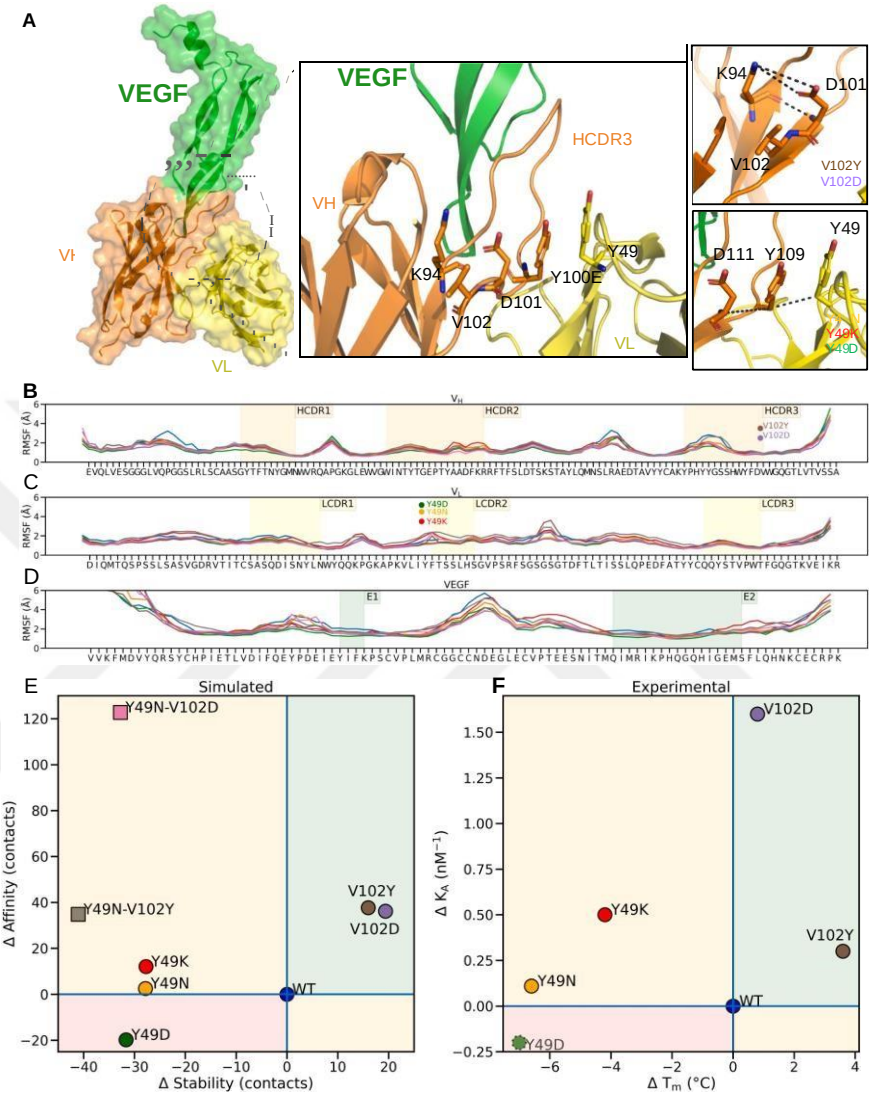
Kabat numbering scheme and the domain definitions were used to determine the complementarity-determining regions (CDRs) and the frameworks (FWs) [58,59] utilizing the web server SabPred-Anarci [60]. Since there are some differences between the Kabat number of the res- idues and the deposited structure number of the residues, the alignments of the numbering schemes were presented in Supplementary Fig. 1. The contacts and distances between V_H and V_L chains were utilized to assess stabilities, and the contacts and distance between both antibody chains and the antigen for affinities. A contact between two interacting do- mains was defined geometrically for when two heavy (*i.e.*, non-hy- drogen) atoms are close to each other within a cutoff of 5 Å or less. Contacts were averaged over all recorded molecular configurations in each trajectory (5000 frames for each simulation). The proximity be- tween two given residues was calculated by the distance between the centers of mass of interacting atoms of given residues. Root mean square fluctuations (RMSFs) were calculated based on the a-carbons of protein chains. The area of the V_H-V_L interface is calculated by subtracting the solvent-accessible surface area (SASA) of the complexed V_H-V_L pair from the sum of the SASAs of the individual V_H and V_L domains. Mass centers, per residue RMSFs, SASA computations were performed by VMD-Python library (https://vmd.robinbetz.com/), and distances between the cen- ters were computed by Python's Numpy. Gromacs version 2018.3 [61] was used for all simulation setups and for the collection of trajectories. VMD [62] and in-house Python scripts were used for all analyses and visualizations.

2.2. In silico secondary structure prediction

The secondary structure of the wild-type and all mutant antibodies were predicted via SABLE prediction webserver [63]. One letter amino acid codes of the antibody sequences were used as input separately. Secondary structure was chosen for prediction goal, SABLE II was cho- sen for server version, WApproximator was chosen for predictor type.

2.3. Protein constructs and protein expression

The anti-VEGF single chain antibody fragment (scFv) heavy chain (V_H) and light chain (V_L) fused via a 21 amino acids length non- repetitive linker "SPNSASHSGSAPQTSSAPGSQ" [53]. The scFv mu- tants were generated by QuikChange Lightning Site-Directed Mutagen- esis Kit (Agilent). The scFv mutants with the leader sequence (PelB), FLAG-tag and polyhistidine-tag were transformed into *E.coli* strain BL21 (DE3) pLysS (Thermo Fisher) with pET17-b (GenScript) expression plasmid. Transformant cells were grown on LB-agar plates containing 100 μg/mL ampicillin and 25 μg/mL chloramphenicol. Single colonies were inoculated in LB broth containing 100 μg/mL ampicillin and 25 μg/ mL chloramphenicol and grown overnight at 225 rpm, 37 ° C. These cells were inoculated into 300 mL autoinduction-media and incubated at 18 °C, 250 rpm for 48 h [64].



(caption on next page)

Fig. 1. Mutations in the vicinity of the conserved salt bridge (V_H-K94:V_H-D101) of the variable heavy chain modulate antibody affinity and stability characteristics. (A) View of the antigen (Vascular Endothelial Growth Factor, VEGF) bound to its cognate scFv fragment (left panel, PBD ID: 1BJ1 [50]). Insets show the salt bridge as well as the nearby mutational landscape investigated in this study (center and right panels, respectively). Orange, yellow, and green colors indicate the scFv heavy chain (V_H), scFv light chain (V_L) and VEGF proteins, respectively. The conserved salt bridge between V_H residues K94 and D101 is indicated via dashed black lines
(top right panel). Rational mutations on scFv involve residues V102 (on V_H) and Y49 (on V_L). (B–D) Molecular dynamics of the wild-type (WT) and mutated scFv constructs. The flexibility of the proteins are illustrated by the average root mean square fluctuations (RMSFs) of backbone atoms of the heavy chain (B), the light chain (C) and the antigen VEGF (D). Blue, orange, red, green, brown, purple, pink and gray lines indicate WT, V_L-Y49N, V_L-Y49K, V_L-Y49D, V_H-V102Y, V_H-V102D, V_L-Y49N-V_H-V102D, V_L-Y49N-V_H-V102Y mutants, respectively. This color convention is used throughout the rest of the text. Complementarity determining regions (CDRs) of the heavy chain, the light chain, and the epitopes of VEGF are highlighted in orange, yellow and green, respectively. Mutated residues are annotated with circles. The time evolution of the V_H:V_L contacts for all mutants are shown in Supplementary Fig. 2. The time evolution of the antibody:antigen contacts for all mutants are shown in Supplementary Fig. 3. (*E*-F) Scatter plots of affinity and stability differentials with respect to the WT antibody. The green areas represent "increased affinity and increased stability", red "decreased affinity and decreased stability" and oranges "decreased affinity or stability", meaning a trade-off. (E) Computed contact differentials from the molecular dynamics trajectories (See Methods). Differences in the mean of total contact counts between the V_H and V_L chains (an indicator of stability, x-axis) and between scFv and VEGF (an indicator of affinity, y-axis) with respect to their WT counterparts are shown. Distribution of contact counts between VH:VL (stability) and antibody antigen (affinity) are shown in Supplementary Fig. 4A and 4B, respectively. (F) Experimentally measured stability

(T_m, thermal melting temperature, x-axis) and affinity (as illustrated via $K_A = 1/K_D$, the association constant, y-axis) differentials of each mutant with respect to the WT. The experimentally unstable construct V_L-Y49D is annotated via a dashed-edged circle. Considering their discouraging MD properties, we did not attempt to produce the two double mutants (V_L-Y49N-V_H-V102D, V_L-Y49N-V_H-V102Y) (square marks in E). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Protein purification

Cultures were centrifuged at 6500 ×*g* and 4 °C (Avanti, Beckman Coulter). Protein containing supernatant was incubated with His-Pur Ni- NTA resin (Thermo Fisher) for 2 h at 4 ° C shaking vigorously. The mixture was loaded into a 10 ml vacuum column (Thermo Fisher) and purified according to recommended commercial protocol. Phosphate buffered saline (PBS) with 25 mM Imidazole, pH 7.4 and PBS with 500 mM Imidazole, pH 7.4 were used as wash and elution buffers, respec- tively. Purified protein was buffer-exchanged into PBS (pH 7.4) through membrane filtration (Amicon® Ultra-4 Centrifugal Filter Units, MWCO 10 kDa, Merck). Protein samples were loaded onto HiTrap ™ Protein L column (GE Healthcare) as a second purification step to maximize protein purity. Protein purities were confirmed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis (TGX ™, FastCast™, 12% Acrylamide kit; Bio-Rad). Precision Plus Protein™ Dual Color standard was used as a marker (Bio-Rad). Protein concentrations were determined by NanoDrop 2000 (at 280 nm). Extinction coefficients were determined 68,550 M^{-1} cm^{-1} for WT and V_H-V102D, 67,060 M^{-1} cm^{-1} for V_L-Y49D, V_L-Y49K, V_L-Y49N, 70,040 M^{-1} cm^{-1} for V_H-V102Y via Expasy ProtParam webserver while using the protein sequences as input [65].

2.5. Thermal denaturation assay

Thermal unfolding profiles of purified scFv proteins were determined by thermal shift assay by ABI 7500 Fast RT-PCR. SYPRO™ Orange Protein Gel Stain (Thermo Fisher) at 5× concentration was used with a 5 μM antibody concentration. Temperature range of 25–99 °C with a 0.05% ramp rate was used. Thermal transition mid-points (*i.e.*, T_m values) from the thermogram data were determined using the Hill equation fit by Origin 8.5 software.

2.6. Surface plasmon resonance (SPR)

Affinity measurements were performed using surface plasmon resonance (SPR) on a Biacore T200 instrument (Biacore Inc., Piscataway, NJ). All experiments were performed in an HBS-EP buffer, pH 7.4. 1000 nM His-tagged VEGF protein was immobilized on a CM4 chip at a flow rate of 10 μl/min for ~1 min (target RU for immobilization was 100 RU). A series of solutions ranging from 10 to 100 nM scFv fragments were subsequently injected at a flow rate of 30 μl/min onto the VEGF- immobilized surface. Regeneration was performed with 10 mM glycine-HCl at pH 2.7 at a flow rate of 30 μl/min for 30 s after each concentration in the run. Data were corrected by double-referencing against a control flow cell containing no VEGF and injecting buffer so- lution. Sensogram curves were analyzed using the BiaEval 3.0 manufacturer's software. K_D, k_{on} and k_{off} values were calculated by fitting the kinetic association and dissociation curves to a 1:1 binding model.

3. Results

3.1. Rationale behind mutational designs

Two positions, one from light chain (V_L-Y49) and one from heavy chain (V_H-V102) are chosen to evaluate affinity/stability trade-offs through secondary interactions of the conserved salt bridge (V_H-K94 and V_H-D101) under the stem of HCDR3. V_L-Y49 is a conserved frame- work residue which is positioned at the V_L-HCDR3 interface. V_L-Y49 makes *n-n* contacts with an HCDR3 residue, V_H-Y100E. On the other hand, V_H-Y100E makes an either hydrophobic or anion-*n* contact with V_H-D101 [66], so there is an anion-*n*-*n* interaction between those three residues (V_H-D101:V_H-Y100E:V_L-Y49, Fig. 1A). For V_L-Y49, we designed three mutations, Y49N to evaluate anion-*n*-amino interaction, Y49K to evaluate anion-*n*-cation interaction, and Y49D, anion-*n*-anion interac- tion. Because this position is surface accessible, all these mutations would also help to improve solubility leading to a possible stability in- crease. However, Y49D mutation is designed as a disruptive mutation due to possibility of repulse between two negatively charged amino acids, V_L-D49:V_H-D101.

V_H-V102 is the last residue of HCDR3, and it does not have contacts with antigen, and it is relatively conserved according to its distribution in *Homo sapiens* (Supplementary Fig. 5). The most common residue is tyrosine (33%) followed by valine (24%). This residue is also high- lighted as one of the key stabilizing contacts for HCDR3 structural di- versity [46]. On the other hand, aspartate in this position is rarely found (1%). We designed two mutations for this position: V_H-V102Y and V_H- V102D. While V_H-V102Y would show the difference between the two most conserved amino acids, V_H-V102D might form complex salt bridge at the stem of HCDR3 and it might also improve stability in soluble conditions due to its negative charge.

3.2. Affinity and stability profiles of designed scFvs

The conserved salt bridge between V_H-K94:V_H-D101 at the stem of HCDR3 defines the robustness of this loop (Supplementary Fig. 5). HCDR3 is the main paratope for most antigens [67], thus we designed several mutations around this salt bridge, preferentially on a Vernier zone residue. We aimed to modulate the antigen affinity without compromising HCDR3 because it is critical that the stem of HCDR3 has a light-chain interface that can contribute to the stability of the antibody. We chose residues V_H-V102 and V_L-Y49 to understand the secondary/ tertiary effects on the V_H-V_L interface (Fig. 1A). In the MD simulations of these constructs, average root mean square fluctuations (RMSF) values near the mutations remained within the range of 6 to 10 Å, indicating that the antibody and antigen structures can still maintain a robust binding configuration. These results suggest that the overall structural flexibility of scFvs is not altered significantly by the introduced muta- tions (Fig. 1B-D). *In silico* secondary structure predictions also showed that point mutations are well tolerated within scFv secondary structure (Supplementary Fig. 6). However, the conserved or drastic mutations on V_H-V102 and V_L-Y49 revealed notable affinity and/or stability changes [68] (Fig. 1B-F, Fig. 2).

On the V_H side, while V_H-V102Y is a mutation to a more conserved residue for this specific position (33% Y, 24% V, Supplementary Fig. 5), V_H-V102D mutation is a drastic change to see whether the ionic bonding of the salt bridge at the stem of HCDR3 is disrupted by the introduction of a proximal acidic residue. Both mutations improve affinity and sta- bility as suggested jointly by our MD simulations and experimental measurements (Fig. 1E, F, Fig. 2). Only V102D has a

higher than pre- dicted affinity while showing slight difference in stability.

HCDR3 is in direct contact with the V_L chain through the V_H-Y100E: V_L-Y49 *n-n* interaction (Fig. 1A). Although V_H-Y100E is in HCDR3, this *n-n* stacking proves to be a crucial contact for the V_H-V_L interface and the HCDR3 robustness through V_H-D101:V_H-Y100E:V_L-Y49 anion-*n*-*n* stacking. We designed V_L-Y49K, V_L-Y49N and V_L-Y49D mutations to convert this triple stacking into anion-*n*-cation, anion-*n*-amino and anion-*n*-anion, respectively. Here, we designed the V_L-Y49D mutant as a negative control to disrupt this stacking interaction and we showed that it has very low stability according to MD simulations. Probably related to this stability loss, it even could not be recombinantly produced (Fig. 1E, F). The other mutants were successfully produced and purified from the supernatant (Fig. 2A). Among the mutants, V_H-V102Y showed the highest increase in thermal stability while V_H-V102D mutant improved the thermal stability slightly compared to WT (Fig. 2C). Those two mutants also showed improved affinity in SPR analysis (Fig. 2B, Supplementary Fig. 7). Although V_L-Y49N, V_L-Y49K mutations showed slightly increased affinity, their stability is compromised according to both computational and experimental findings (Fig. 1B-F, Fig. 2B-C, Supplementary Fig. 7).

3.3. Rescue mutations have non-additive effects

To restore the stability loss of V_L-Y49N by the stability-favoring mutations (V_H-V102D and V_H-V102Y), we designed two *in silico* dou- ble mutants, V_L-Y49N:V_H-V102D and V_L-Y49N:V_H-V102Y. Surprisingly, both rescue mutants showed even worse stability profiles as observed in the MD simulations (Fig. 1E). As a result, we did not pursue these mu- tants experimentally. Most importantly, this finding shows that the ef- fects of single point mutations are not additive, and more complex secondary/tertiary interactions are at play for affinity/stability profiles. We tried to have an insight into the affinity/stability profiles of our designed mutants through more detailed structural studies.

3.4. HCDR3 is essential for both affinity and stability

Computed pairwise contact counts between the protein components of V_L-Y49N suggested specific regions for the increase in affinity and the loss in stability (Fig. 3). While the affinity gain is mostly due to HCDR3- VEGF contact increase as expected, stability decrease is mostly caused by V_H-V_L interface disruption through HCDR3, LCDR3 and/or light framework 2 (LFW2) contact loss (Fig. 3A, B). As expected, VEGF binding occurs mostly with V_H chain (Fig. 3C). Same regions (HCDR3- VEGF for affinity, HCDR3, LCDR3 and LFW2 for stability) play similar roles in other mutants (Supplementary Figs. 8–13). Although V_H-V102Y and V_H-V102D mutants are not directly in the V_H-V_L interface, they have a drastic stability increase probably due to secondary HCDR3 in- teractions and/or stronger intra-HCDR3 contacts. We can examine the HCDR3 loop on two sides, one face to VEGF (HCDR3 residues 95-100B) that plays roles in binding and shows mostly change in VEGF binding, while the other face (HCDR3 residues 100B-102) shows changes in V_L interacting surface (Fig. 1A). This emphasizes that HCDR3 has crucial roles in both affinity and stability profiles of antibodies.

3.5. V_H-V_L interface is compromised in low-stability mutants

The V_H-V_L interface and its packing are known to have a significant effect on the stability of an antibody [69,70]. We analyzed the V_H-V_L interface by computing the buried surface areas (BSAs) of each mutant (Fig. 4, Supplementary Fig. 14). We calculated this property by sub- tracting solvent accessible surface area (SASA) of the complete V_H-V_L complexes from the sum of the SASAs of the individual V_H/V_L proteins (Fig. 4A). Mutants with high measured thermal melting points (V_H- V102Y, V_H-V102D) also have computed BSA values close to or higher than that of the WT (~1750 Å^2 or higher). On the other hand, mutants

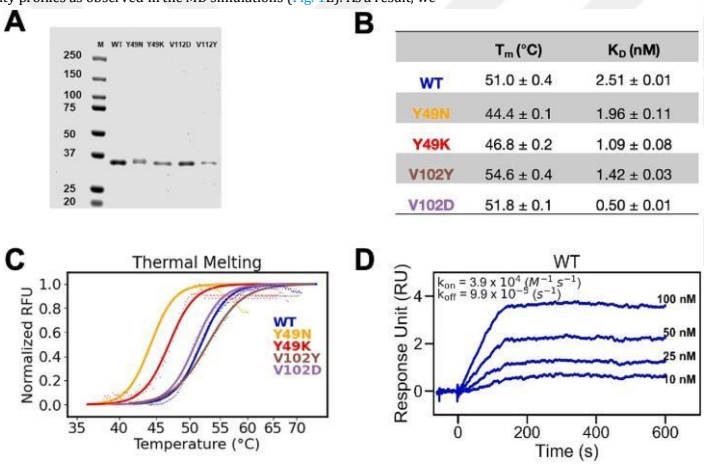


Fig. 2. Experimental affinity and stability profiles of scFv constructs in this study. (A) SDS-PAGE analysis of scFvs after bacterial expression and purification, (B) Thermal melting temperature (T_m, in degrees) and dissociation constant (K_D, in molarity units) values, (C) T_m plots (repeated at least two times in three replicates with different batches of protein) (D) SPR profile of WT for affinity determination (SPR profiles of all mutants are in Supplementary Fig. 7). Color coding follows Fig. 1.

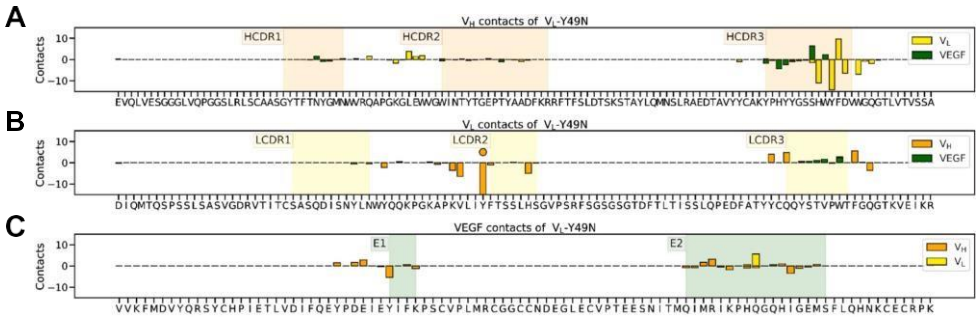


Fig. 3. Difference in pairwise contact counts of V_L-Y49N mutant with respect to WT. V_H, V_L and VEGF interactions per residue are colored orange, yellow and green, respectively. (A) V_L and VEGF contact count differences for V_H residues. (B) V_H and VEGF contact count differences for V_L residues. Mutated residue is annotated with a circle. (C) V_H and V_L contact count differences for VEGF residues. FW regions are not shown: FW1 is before CDR1, FW2 is in between CDR1-CDR2, FW3 is in between CDR2-CDR3, FW4 is in after CDR3. Color coding follows Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

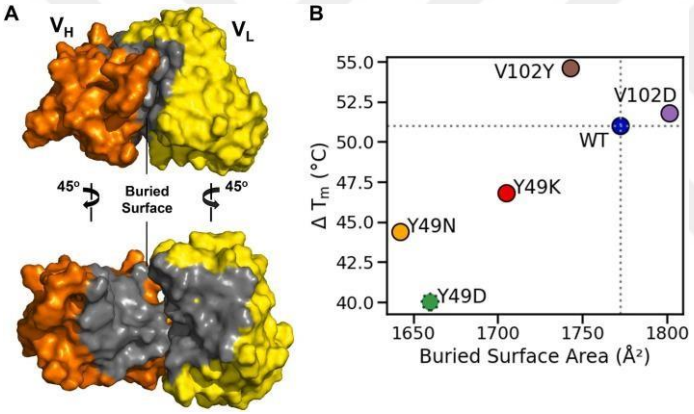


Fig. 4. Correlation between measured thermal stabilities and computed V_H-V_L buried surface areas of all mutants. The area of the V_H-V_L interface is calculated by subtracting the solvent- accessible surface area (SASA) of the complexed V_H-V_L pair from the sum of the SASAs of the in- dividual V_H and V_L domains. (A) Surface repre- sentation of the calculated buried surface area. Surfaces of V_H, V_L and their buried surface areas are colored orange, yellow and gray, respec- tively. (B) Computed average buried surface areas are scattered with the experimental melting temperature values (T_m). The low-stability mutant V_L-Y49D (annotated with a dashed cir- cle) is arbitrarily assigned a melting temperature of 40 degrees. Color coding follows Fig. 1. Dis- tributions of buried surface area of all mutants are shown in Supplementary Fig. 14. (For inter- pretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with low measured thermal melting points (V_L-Y49N, V_L-Y49K, V_L-Y49D) also have computed BSA values that are less than of that of the WT (lower than 1700 Å², Fig. 4B).

3.6. Interaction between V_L-Y49 and HCDR3 V_H-Y100 is a critical determinant of stability

We examined the anion-T-T interactions between the triplet V_H-D101:V_H-Y100E:V_L-Y49 in all mutants (Fig. 5A, B). The distances be- tween the alpha carbons of these residues showed that stability- improved mutants (V_H-V102Y, V_H-V102D) mimic those values remi- niscent to the WT data while stability-compromised mutants have dis- rupted the interactions particularly between V_L-Y49 and V_H-Y100E (Fig. 5C). The V_L-Y49:V_H-100E interaction is a T-T stacking contact that is located at the HCDR3-LFW2 interface, so we checked whether other interactions on this V_H-V_L interface have a role in this stability loss, but no significant relationship wass found (Supplementary Figs. 15, 16). This result shows that HCDR3 has a substantial effect on stability, especially through the residue V_H-Y100E. Even the addition of affinity/stability increasing V_H-V102D/Y mutations did not rescue the stability of V_L-Y49N mutant, it got even worse (Fig. 1E). Therefore, V_L-Y49:V_H- Y100E interaction at the core of the V_H-V_L interface is proven to be very crucial for the overall antibody stability.

3.7. A de novo salt bridge near the stem of HCDR3 leads to a substantial affinity improvement in V102D

While the improvements in stability in the two V_H-V102 mutants can be primarily traced to a more robust HCDR3-V_L interface, affinity im- provements occur mainly through the improved HCDR3-VEGF in- teractions (Supplementary Figs. 10, 11). The V_H-V102D forms a complex salt bridge where one residue forms ionic interactions with more than one residue (Fig. 5B) [48,71]. The core salt bridge (V_H-K94:V_H-D101) is accompanied by a mutated aspartate (V_H-K94:V_H-D102). Although the backbone ionic interaction of V_H-K94:V_H-D101 is not disrupted at all, side-chain ionic interactions of V_H-K94 were shared with V_H-D101 and V_H-D102 for both V_H-V102D and V_L-Y49N-V_H-V102D mutants (Fig. 5D). Forming this complex salt bridge between V_H-

K94:V_H-D101/D102 might contribute to the affinity increase for both V_H-V102D and V_L- Y49N-V_H-V102D mutants. The indirect effect of V_H-V102 residue on

6

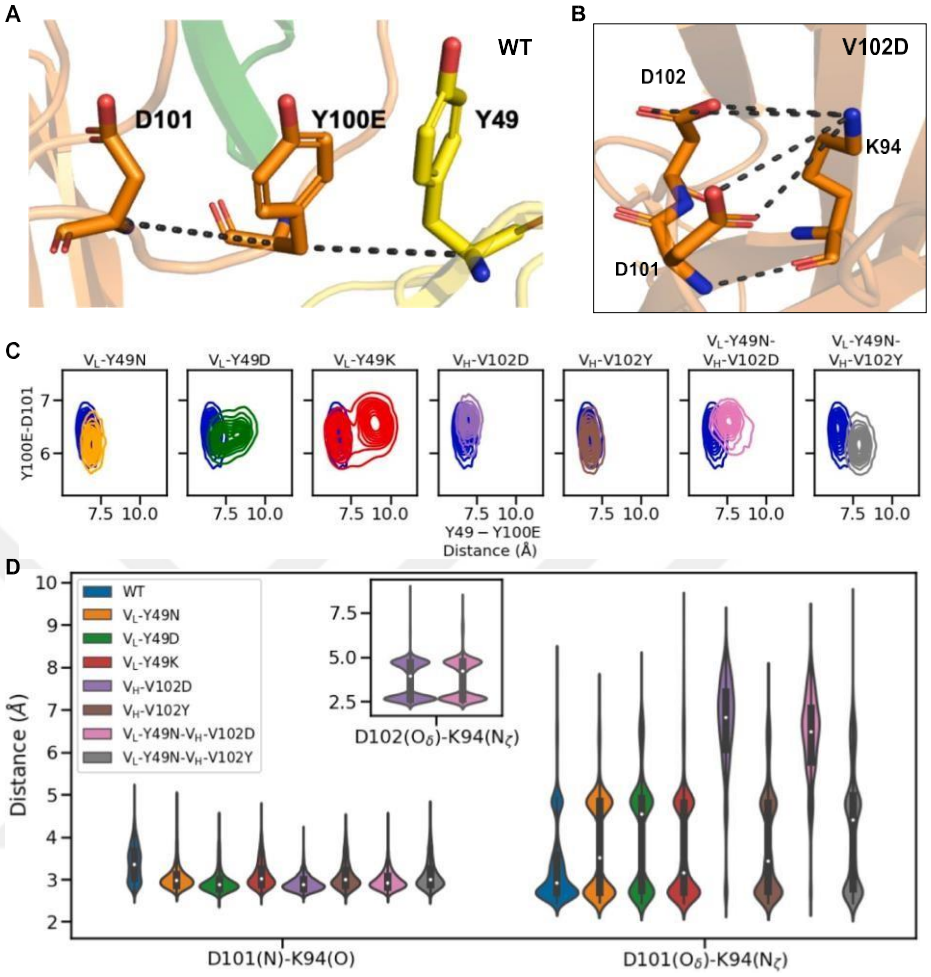


Fig. 5. Antibody stability as assessed via critical inter-residue distances at the V_H-V_L interface. (A) The anion-re-re interaction (V_H-D101:V_H-Y100E:V_L-Y49) at the V_H- V_L interface of HCDR3-LFW2. (B) The core salt bridge V_H-K94:V_H-D101 and the complex salt bridge V_H-K94: V_H-D101/D102 in V_H-V102D (C) Distances between the C_α atoms of V_L-Y49:V_H-Y100E and V_H-Y100E:V_H-D101 residues. Distance values of mutants are plotted together with the WT counterpart to demonstrate the shifts.
(D) Distances between the O atom of V_H-K94 and N atom of V_H-D101 that is the backbone salt bridge and distances between the N_s atom of V_H-K94 and O_8 atom of V_H-D101 that is the side-chain salt bridge for all variants. V_H-V102D mutation is also invented a new ionic interaction with its side chain oxygen atom that competes with the ionic interaction of V_H-D101 oxygen atom, shown as an inset. Color coding follows Fig. 1.

both affinity and stability through HCDR3 loop conformation is notable and worth investigating.

3.8. The complex salt bridge at the stem of HCDR3 improves packing of epitope-paratope interaction

When we checked epitope-paratope interactions, VEGF-I91 is found to be the key player by having an interaction with V_H-H101 of HCDR3 (Supplementary Fig. 17). We recognized that V_H-H101:VEGF-I91 inter- action might explain drastic affinity changes for affinity improved

mutants (V_H-V102D, V_H-V102Y, V_L-Y49N-V_H-V102D, V_L-Y49N-V_H- V102Y) by having less distant interaction overall (Supplementary Fig. 17). There are also other important epitope-paratope interactions such as V_H-Y102:VEGF-I80, V_H-Y102:VEGF-G92, V_H-G100:VEGF-R82 (HCDR3-VEGF for all), but no significant difference was observed except for double mutants which have the highest improvements in their VEGF affinities that can be attributed to *de novo* contacts formed between side chains and backbone functional groups (Supplementary Fig. 18, 19).

To see whether the packing of HCDR3 with VEGF has any contri- bution to observed affinity, we analyzed the HCDR3 conformation

7

change for V_H-V102D (Fig. 6A). An angle is calculated to represent tilt of

HCDR3 towards VEGF (Fig. 6B). To measure this angle, two vectors, $\overrightarrow{P_1}$ between the center of masses of scFv and VEGF, P_2 between the center of masses scFv and middle residues of HCDR3 (V_H-G100 $C\alpha$ and V_H-S100A $C\alpha$) are used. θ angle is determined as the cosine angle between those $\overrightarrow{P_1}$ and $\overrightarrow{P_2}$ vectors (Fig. 6B, Supplementary Fig. 20). Scatter plots of calculated θ angle and experimental affinity change showed that there is an obvious correlation for HCDR3 conformation with affinity changes (Fig. 6C). While mutants with slight affinity increase (V_L-Y49N, V_L- Y49K, V_H-V102Y) are clustered together with angles values very close to that of WT, mutant with the most

significant affinity change (V_H-V102D) had more increase from those of all mutants (Fig. 6C). This might show that the formation of a complex salt bridge at the stem of HCDR3 might alter overall HCDR3 conformation tilting towards VEGF affecting its affinity drastically.

4. Discussion

Monoclonal antibodies are promising biomacromolecules for various therapeutic and diagnostic applications. However, numerous trade-offs can be encountered during development and improvement stages due to the intrinsic complexity and the structural limited modularity of these molecules [72]. While the antigen affinity is the most natural and crucial developability parameter, its improvement cannot be decoupled from the stability of the antibody in a systematic manner, resulting in an unavoidable affinity *versus* stability trade-off [11,15,73]. A primary driver of this mutual dependence could lie in the architecture of the HCDR3 region, an elongated loop that forms

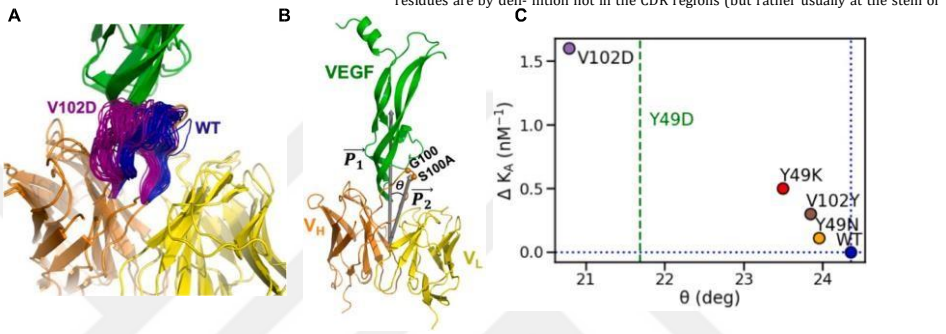


Fig. 6. Changes in affinity correlate with a global tilt in HCDR3. (A) HCDR3 loop conformations are visualized for WT and V_H-V102D. Alignment is performed based on the reference of the V_H-K94:V_H-D101 salt bridge. HCDR3s are represented as the tubes,----- HCDR3- ------- frames ---- of V_H-V102D and WT are colored in purple and blue, respectively. (B) The cosine angle (*θ*) change of HCDR3 is calculated between two vectors and vector is the vector between the center of mass of the scFv fragment and the center of mass of antigen VEGF. ----- P_1 P_2 P_1 is the vector between the center of mass of the scFv fragment and the center of masses of V -G100 and V -S100A C atoms. and---- P_2 vectors are colored in gray. Selected atoms are represented as spheres. (C) Angle decrease of mutants compared to WT-HCDR3 are $_o P_1 P_2$ scattered through every 100th frame of the whole trajectory (5000 frames). The experimental affinity of the V_L-Y49D mutant could not be obtained, so its repre- sentation is made based on our computational results. Color coding follows Fig. 1. Distributions of angle of all mutants are shown in Supplementary Fig. 20. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

There is a highly conserved salt bridge (V_H-K94:V_H-D101) at the stem of HCDR3 that is critical for HCDR3 to assume its bulge form [85]. If the critical interactions with both the antigen and the antibody light chain.

In this study, we focused on two pivotal residues (Y49 on the light chain and V102 on the heavy chain) that modulate the global orientation of HCDR3 while maintaining its overall shape and structural integrity (Fig. 1). Of these two residues, the light chain Y49 tolerated a mutation that improved the antigen affinity but a crucial *n-n* interaction was lost in the V_H-V_L interface, one that proved irreplaceable for the stability. The second residue of focus, heavy chain V102, not only tolerated mutations that reoriented HCDR3 more favorably for antigen binding, but interactions lost with the light chain could be compensated via novel contacts not present in WT. We identified one such mutation, V_H-V102Y, which demonstrated a joint increase in stability and affinity in both our experimental measurements and atomistic molecular dy- namics simulations. HCDR3 has two main interfaces, one towards the antigen and the other one spanning part of the V_H-V_L interface. We have shown that in these mutants (i) the contacts at the HCDR3-VEGF inter- face increase, resulting in an improved affinity and (ii) the contacts at the HCDR3, LFW2, and LCDR3 region of the V_H-V_L interface increase, probably resulting in an improved stability. These two mutants are direct evidence that HCDR3 is not only important for affinity but also for stability and not necessarily in an antagonistic manner. We note that caution should be taken to avoid drastic trade-offs during affinity maturation efforts on HCDR3. The findings here are a novel proof of concept demonstration with implications for other antibody engineering efforts.

Understanding antibody stability is a more complex issue because the residues contributing to stability are scattered across diverse posi- tions (core domain, surface exposed residues, V_H-V_L interface) [74–77]. In our study, stability increasing mutations (V_H-V102Y and V_H-V102D) showed significantly higher V_H-V_L interface buried surface area (Fig. 4). This demonstrates the importance of the scope of V_H-V_L packing for the overall antibody stability. In addition, the V_H-V_L orientation is also known to be important for stability [78–81]. In our mutants, no signif- icant correlation was observed between this metric and stability.

It is known that *n-n* stacking is very important for protein structure and function. Besides *n-n*, *n*-cation/amino/anion contacts are also a part of these crucial stacking interactions [82]. Even triple *n*-stackings are known to contribute to the activity of proteins [83]. In the specific example of light chain Y49, we mutated this residue to collect structural insights into the effect of triple *n*-stacking interactions at the V_H-V_L interface. We showed that drastic affinity and stability changes occur when the anion-*n*-*n* interaction (V_H-D101: V_H-Y100E:V_L-Y49) was mutated through V_L-Y49N/K/D mutations (Fig. 4, 5). V_L-Y49 is one of the Vernier zone residues. Although Vernier zone residues are by defi- nition not in the CDR regions (but rather usually at the stem of CDRs), they are known to be indispensable for antibody function [24]. These residues are usually back-mutated to restore antibody affinity in humanization efforts [84]. Here, we provide further evidence on the importance of Vernier zone residues for antibody engineering efforts [27].

lysine (or arginine) at position 94 is converted to any other amino acid, HCDR3 loses function due to lack of stabilizing salt bridge and it does not form the bulge [86]. When this salt bridge was converted into a complex salt bridge by V_H-V102D mutation, affinity and stability increased. In- depth analysis of a molecular dynamics simulation of this mutant showed that the complex salt bridge at the stem shifted the HCDR3 loop to tilt towards VEGF, thereby contributing to the affinity increase (Fig. 6). Even though V_H-V102 is a highly conserved residue in the antibody framework [87], it is nonetheless a potentially interesting locus for future antibody engineering and affinity maturation efforts. As a general approach, V102 might be mutated to "D" or "Y" to increase antibody affinity with no loss or even better stability. Aspartate in this position is very rare (1%), thus immunogenicity should also be consid- ered while designing mutations with rare amino acids in particular positions. We should also note here the complementarity aspect of our molecular dynamics simulations in understanding the detailed molecular mechanisms associated with the involvement of HCDR3. Because all the domains of an antibody share the same structural fold [19,88] and framework residues, especially chosen HCDR3 salt bridge interactions which are usually very conserved among organisms, our findings can be applied to other antibodies in general. The measured energetic differ- ences in the affinities are on the order of a few thermal energies at ambient conditions (as inferred from the dissociation constants), meaning that accurate

Acknowledgments

prediction of protein-protein binding energies is of utmost importance in this context. As an intuitive measure of the relative changes in binding interfaces, here, we have used pairwise contacts between the epitope and paratope groups. Such geometric metrics are commonly employed in heuristic correlations with experimentally measured energies [89].

Affinity improvement or re-gaining efforts are usually encountered with numerous trade-offs such as loss of stability, lower solubility, and/or higher aggregation propensity, as reviewed elsewhere [11]. Although these *in vitro* properties of a candidate antibody can be co-screened with a variety of experimental tools, scalability is typically costly and cumbersome. In this context, the use of *in silico* tools can alleviate the load by providing precise predictions at a fraction of the cost and time typically invested in an experimental undertaking. In this work, we tapped into the strength of sufficiently long molecular dynamics simu- lations which not only validated our physico-chemical wet-lab charac- terization of our mutants but also provided molecular level understanding into the favorable and unfavorable outcomes.

Declaration of Competing 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.

Data availability

Data will be made available on request.

Author contributions

MA, TU, S.Kale and S.Kalyoncu contributed to conception and design of the study. MA carried out *in vitro* experiments. TU and MA carried out *in silico* analysis. TU and MA performed data visualization. MA and TU wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version. S.Kale and S.Kalyoncu supervised the project.

This work was supported by the institutional funds of Izmir Biomedicine and Genome Center (IBG) (T.U., S.Kale, and S.Kalyoncu), Council of Higher Education (YÖK) 100/2000 Fellowship Program (M. A.), The Scientific and Technological Research Council of Turkey (TUBITAK)-BIDEB 2211A PhD Achievement Grant (M.A.), TUBITAK grant 119Z161 (M.A., S.Kale, and S.Kalyoncu), and EMBO installation grant no. 5056 (S.Kale). The authors acknowledge the computational resources of TUBITAK ULAKBIM High Performance and Grid Computing Center (TRUBA resources) and the IBG High Performance Computing Facility.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbapap.2023.140915.

References

[1] H. Kaplon, M. Muralidharan, Z. Schneider, J.M. Reichert, Antibodies to watch in 2020, MAbs. 12 (1) (2020) 1703531.
[2] H. Kaplon, J.M. Reichert, Antibodies to watch in 2021, MAbs. 13 (1) (2021) 1860476.
[3] A. Bates, C.A. Power, David vs., Goliath: the structure, function, and clinical prospects of antibody fragments, Antibodies (Basel) 8 (2) (2019).
[4] G. Chao, W.L. Lau, B.J. Hackel, S.L. Sazinsky, S.M. Lippow, K.D. Wittrup, Isolating and engineering human antibodies using yeast surface display, Nat Protoc 1 (2) (2006) 755–768.
[5] D.J. Schofield, A.R. Pope, V. Clementel, J. Buckell, S. Chapple, K.F. Clarke, et al., Application of phage display to high throughput antibody generation and characterization, Genome Biol 8 (11) (2007) R254.
[6] J. Santos, M. Cardoso, I.S. Moreira, J. Goncalves, J.D.G. Correia, S.C. Verde, et al., Integrated in silico and experimental approach towards the Design of a Novel Recombinant Protein Containing an anti-HER2 scFv, Int J Mol Sci 22 (7) (2021).
[7] A. Goel, D. Colcher, J. Baranowska-Kortylewicz, S. Augustine, B.J. Booth, G. Pavlinkova, et al., Genetically engineered tetravalent single-chain Fv of the pancarcinoma monoclonal antibody CC49: improved biodistribution and potential for therapeutic application, Cancer Res 60 (24) (2000) 6964–6971.
[8] B.J. Fennell, B. McDonnell, A.S. Tam, L. Chang, J. Steven, I.D. Broadbent, et al., CDR-restricted engineering of native human scFvs creates highly stable and soluble bifunctional antibodies for subcutaneous delivery, MAbs. 5 (6) (2013) 882–895.
[9] M.S. Sawant, C.N. Streu, L. Wu, P.M. Tessier, Toward drug-like multispecific antibodies by design, Int J Mol Sci 21 (20) (2020).
[10] M.C. Julian, C.C. Lee, K.E. Tiller, L.A. Rabia, E.K. Day, A.J. Schick 3rd, et al., Co- evolution of affinity and stability of grafted amyloid-motif domain antibodies, Protein Eng Des Sel 28 (10) (2015) 339–350.
[11] L.A. Rabia, A.A. Desai, H.S. Jhajj, P.M. Tessier, Understanding and overcoming trade-offs between antibody affinity, specificity, stability and solubility, Biochem Eng J 137 (2018) 365–374.
[12] G. Houlihan, P. Gatti-Lafranconi, D. Lowe, F. Hollfelder, Directed evolution of anti- HER2 DARPins by SNAP display reveals stability/function trade-offs in the selection process, Protein Eng Des Sel 28 (9) (2015) 269–279.
[13] M.S. Faber, T.A. Whitehead, Data-driven engineering of protein therapeutics, Curr Opin Biotechnol 60 (2019) 104–110.
[14] R. Cohen-Khait, O. Dym, S. Hamer-Rogotner, G. Schreiber, Promiscuous protein binding as a function U S A 114 (4) (2017) E486–E495.
[17] J.L. Xu, M.M. Davis, Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities, Immunity. 13 (1) (2000) 37–45.
[18] J. Bostrom, C.V. Lee, L. Haber, G. Fuh, Improving antibody binding affinity and specificity for therapeutic development, Methods Mol Biol 525 (2009) 353–376 (xiii).
[19] M.L. Chiu, D.R. Goulet, A. Teplyakov, G.L. Gilliland, Antibody structure and function: the basis for engineering therapeutics, Antibodies (Basel) 8 (4) (2019).
[20] E.M. Herold, C. John, B. Weber, S. Kremser, J. Eras, C. Berner, et al., Determinants of the assembly and function of antibody variable domains, Sci Rep 7 (1) (2017) 12276J. Zhao,. R. Nussinov, B. Ma, The allosteric effect in antibody-antigen recognition,
[21] Methods Mol Biol 2253 (2021) 175–183.
[22] K. Makabe, T. Nakanishi, K. Tsumoto, Y. Tanaka, H. Kondo, M. Umetsu, et al., Thermodynamic consequences of mutations in Vernier zone residues of a humanized anti-human epidermal growth factor receptor murine antibody, 528, J Biol Chem 283 (2) (2008) 1156–1166.

of protein stability, Structure. 25 (12) (2017) 1867–74 e3.
[15] M.C. Julian, L. Li, S. Garde, R. Wilen, P.M. Tessier, Efficient affinity maturation of antibody variable domains requires co-selection of compensatory mutations to maintain thermodynamic stability, Sci Rep 7 (2017) 45259.
[16] P. Koenig, C.V. Lee, B.T. Walters, V. Janakiraman, J. Stinson, T.W. Patapoff, et al., Mutational landscape of antibody variable domains reveals a switch modulating the interdomain conformational dynamics and antigen binding, Proc Natl Acad Sci
[23] L. Riechmann, M. Clark, H. Waldmann, G. Winter, Reshaping human antibodies for therapy, Nature. 332 (6162) (1988) 323–327.
[24] J. Foote, G. Winter, Antibody framework residues affecting the conformation of the hypervariable loops, J Mol Biol 224 (2) (1992) 487–499.
[25] H. Wu, Y. Nie, W.D. Huse, J.D. Watkins, Humanization of a murine monoclonal antibody by simultaneous optimization of framework and CDR residues, J Mol Biol 294 (1) (1999) 151–162.
[26] Y. Safdari, S. Farajnia, M. Asgharzadeh, M. Khalili, Antibody humanization methods - a review and update, Biotechnol Genet Eng Rev 29 (2013) 175–186.
[27] M. Arslan, D. Karadag, S. Kalyoncu, Conformational changes in a Vernier zone region: implications for antibody dual specificity, Proteins. 88 (11) (2020) 1447–1457.
[28] M. Arslan, D. Karadag, S. Kalyoncu, Protein engineering approaches for antibody fragments: directed evolution and rational design approaches, Turk J Biol 43 (1) (2019) 1–12.
[29] K.E. Tiller, P.M. Tessier, Advances in antibody design, Annu Rev Biomed Eng 17 (2015) 191–216.
[30] J. Zhao, R. Nussinov, W.J. Wu, B. Ma, In silico methods in antibody design, Antibodies (Basel) 7 (3) (2018).
[31] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, Molecular dynamics study of unbinding of the avidin-biotin complex, Biophys J 72 (4) (1997) 1568–1581.
[32] D. Dogan, M. Arslan, T. Ulucay, S. Kalyoncu, S. Dimitrov, S. Kale, CENP-A nucleosome is a sensitive allosteric scaffold for DNA and chromatin factors, J Mol Biol 433 (6) (2021), 166789.
[33] A.G. Rocco, L. Mollica, P. Ricchiuto, A.M. Baptista, E. Gianazza, I. Eberini, Characterization of the protein unfolding processes induced by urea and temperature, Biophys J 94 (6) (2008) 2241–2251.
[34] K. Lindorff-Larsen, N. Trbovic, P. Maragakis, S. Piana, D.E. Shaw, Structure and dynamics of an unfolded protein examined by molecular dynamics simulation, J Am Chem Soc 134 (8) (2012) 3787–3791.
[35] F. Collu, E. Spiga, N. Chakroun, H. Rezaei, F. Fraternali, Probing the early stages of prion protein (PrP) aggregation with atomistic molecular dynamics simulations, Chem Commun (Camb) 54 (57) (2018) 8007–8010.
[36] M.L. Fernandez-Quintero, N.D. Pomarici, B.A. Math, K.B. Kroell, F. Waibl, A. Bujotzek, et al., Antibodies exhibit multiple paratope states influencing VH-VL domain orientations, Commun Biol 3 (1) (2020) 589.
[37] T. Yamashita, Toward rational antibody design: recent advancements in molecular dynamics simulations, Int Immunol 30 (4) (2018) 133–140.
[38] X. Cheng, J. Wang, G. Kang, M. Hu, B. Yuan, Y. Zhang, et al., Homology modeling-based in silico affinity maturation improves the affinity of a Nanobody, Int J Mol Sci 20 (17) (2019).
[39] B. Yang, S.J. Lin, J.Y. Ren, T. Liu, Y.M. Wang, C.M. Li, et al., Molecular docking and molecular dynamics (MD) simulation of human anti-complement factor H (CFH) antibody Ab42 and CFH polypeptide, Int J Mol Sci 20 (10) (2019).
[40] K. Yoshida, D. Kuroda, M. Kiyoshi, M. Nakakido, S. Nagatoishi, S. Soga, et al., Exploring designability of electrostatic complementarity at an antigen-antibody interface directed by mutagenesis, biophysical analysis, and molecular dynamics simulations, Sci Rep 9 (1) (2019) 4482.
[41] S. Chiba, A. Tanabe, M. Nakakido, Y. Okuno, K. Tsumoto, M. Ohta, Structure-based design and discovery of novel anti-tissue factor antibodies with cooperative double-point mutations, using interaction analysis, Sci Rep 10 (1) (2020) 17590.
[42] Y.N. Zhang, X.Q. Zhang, X.C. Zhang, J.W. Xu, L.L. Li, X.Y. Zhu, et al., Key amino acid residues influencing binding affinities of pheromone-binding protein from Athetis lepigone to two sex pheromones, J Agric Food Chem 68 (22) (2020) 6092–6103.
[43] M.L. Fernandez-Quintero, K.B. Kroell, F. Hofer, J.R. Riccabona, K.R. Liedl, Mutation of framework residue H71 results in different antibody Paratope states in solution, Front Immunol 12 (2021), 630034.
[44] B. Musafia, V. Buchner, D. Arad, Complex salt bridges in proteins: statistical analysis of structure and function, J Mol Biol 254 (4) (1995) 761–770.
[45] H. Shirai, A. Kidera, H. Nakamura, H3-rules: identification of CDR-H3 structures in antibodies, FEBS Lett 455 (1–2) (1999) 188–197.
[46] B.D. Weitzner, R.L. Dunbrack Jr., J.J. Gray, The origin of CDR H3 structural diversity, Structure. 23 (2) (2015) 302–311.
[47] K. Tsumoto, K. Ogasahara, Y. Ueda, K. Watanabe, K. Yutani, I. Kumagai, Role of salt bridge formation in antigen-antibody interaction. Entropic contribution to the complex between hen egg white lysozyme and its monoclonal antibody HyHEL10, J Biol Chem 271 (51) (1996) 32612–32616.
[48] L.G. Presta, H. Chen, S.J. O'Connor, V. Chisholm, Y.G. Meng, L. Krummen, et al., Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders, Cancer Res 57 (20) (1997) 4593–4599.
[49] K. Zirlik, J. Duyster, Anti-Angiogenics: current situation and future perspectives, Oncol Res Treat 41 (4) (2018) 166–171.
[50] Y.A. Muller, Y. Chen, H.W. Christinger, B. Li, B.C. Cunningham, H.B. Lowman, et al., VEGF and the fab fragment of a humanized neutralizing antibody:

crystal structure of the complex at 2.4 a resolution and mutational analysis of the interface, Structure. 6 (9) (1998) 1153–1167.
[51] P.J. Hudson, A.A. Kortt, High avidity scFv multimers; diabodies and triabodies, J Immunol Methods 231 (1–2) (1999) 177–189.
[52] F. Le Gall, S.M. Kipriyanov, G. Moldenhauer, M. Little, Di-, tri- and tetrameric single chain Fv antibody fragments against human CD19: effect of valency on cell binding, FEBS Lett 453 (1–2) (1999) 164–168.
[53] M. Arslan, M. Karadag, E. Onal, E. Gelinci, G. Cakan-Akdogan, S. Kalyoncu, Effect of non-repetitive linker on in vitro and in vivo properties of an anti-VEGF scFv, Sci Rep 12 (1) (2022) 5449.
[54] M.L. Fernandez-Quintero, V.J. Hoerschinger, L.M. Lamp, A. Bujotzek, G. Georges, K.R. Liedl, VH -VL interdomain dynamics observed by computer simulations and NMR, Proteins. 88 (7) (2020) 830–839.
[55] R.B. Best, X. Zhu, J. Shim, P.E. Lopes, J. Mittal, M. Feig, et al., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles, J Chem Theory Comput 8 (9) (2012) 3257–3273.
[56] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, et al., CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat Methods 14 (1) (2017) 71–73.
[57] S. Izadi, R. Anandakrishnan, A.V. Onufriev, Building water models: a different approach, J Phys Chem Lett 5 (21) (2014) 3863–3871.
[58] T.T. Wu, E.A. Kabat, An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity, J Exp Med 132 (2) (1970) 211–250.
[59] K.R. Abhinandan, A.C. Martin, Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains, Mol Immunol 45 (14) (2008) 3832–3839.
[60] J. Dunbar, C.M. Deane, ANARCI: antigen receptor numbering and receptor classification, Bioinformatics. 32 (2) (2016) 298–300.
[61] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen, GROMACS: fast, flexible, and free, J Comput Chem 26 (16) (2005) 1701–1718.
[62] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J Mol Graph 14 (1) (1996), 33–8, 27–8.
[63] R. Adamczak, A. Porollo, J. Meller, Combining prediction of secondary structure and solvent accessibility in proteins, Proteins. 59 (3) (2005) 467–475.
[64] F.W. Studier, Protein production by auto-induction in high density shaking cultures, Protein Expr Purif 41 (1) (2005) 207–234.
[65] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.C. Sanchez, K.L. Williams, R.D. Appel, et al., Protein identification and analysis tools in the ExPASy server, Methods Mol Biol 112 (1999) 531–552.
[66] X. Lucas, A. Bauza, A. Frontera, D. Quinonero, A thorough anion-pi interaction study in biomolecules: on the importance of cooperativity effects, Chem Sci 7 (2) (2016) 1038–1050.
[67] S. D'Angelo, F. Ferrara, L. Naranjo, M.F. Erasmus, P. Hraber, A.R.M. Bradbury, Many routes to an antibody heavy-chain CDR3: necessary, yet insufficient, for specific binding, Front Immunol 9 (2018) 395.
[68] M.B. Swindells, C.T. Porter, M. Couch, J. Hurst, K.R. Abhinandan, J.H. Nielsen, et al., abYsis: integrated antibody sequence and structure-management, analysis, and prediction, J Mol Biol 429 (3) (2017) 356–364.
[69] P.H. Tan, B.M. Sandmaier, P.S. Stayton, Contributions of a highly conserved VH/ VL hydrogen bonding interaction to scFv folding stability and refolding efficiency, Biophys J 75 (3) (1998) 1473–1482.
[70] K. Masuda, K. Sakamoto, M. Kojima, T. Aburatani, T. Ueda, H. Ueda, The role of interface framework residues in determining antibody V(H)/V(L) interaction strength and antigen-binding affinity, FEBS J 273 (10) (2006) 2184–2194.
[71] A.G. Gvritishvili, A.V. Gribenko, G.I. Makhatadze, Cooperativity of complex salt bridges, Protein Sci 17 (7) (2008) 1285–1290.
[72] L.S. Bigman, Y. Levy, Proteins: molecules defined by their trade-offs, Curr Opin Struct Biol 60 (2020) 50–56.
[73] L. Shehata, D.P. Maurer, A.Z. Wec, A. Lilov, E. Champney, T. Sun, et al., Affinity maturation enhances antibody specificity but compromises conformational stability, Cell Rep 28 (13) (2019) 3300–8 e4.
[74] A. Worn, A. Pluckthun, Different equilibrium stability behavior of ScFv fragments:
identification, classification, and improvement by protein engineering, Biochemistry. 38 (27) (1999) 8739–8750.
[75] H.J. Chang, J.W. Jian, H.J. Hsu, Y.C. Lee, H.S. Chen, J.J. You, et al., Loop-sequence features and stability determinants in antibody variable domains by high-throughput experiments, Structure. 22 (1) (2014) 9–21.
[76] A. Lehmann, J.H. Wixted, M.V. Shapovalov, H. Roder, R.L. Dunbrack Jr., M.
K. Robinson, Stability engineering of anti-EGFR scFv antibodies by rational design of a lambda-to-kappa swap of the VL framework using a structure-guided approach, MAbs. 7 (6) (2015) 1058–1071.
[77] L. Montoliu-Gaya, J.C. Martinez, S. Villegas, Understanding the contribution of disulfide bridges to the folding and misfolding of an anti-Abeta scFv, Protein Sci 26 (6) (2017) 1138–1149.
[78] A. Worn, A. Pluckthun, Mutual stabilization of VL and VH in single-chain antibody fragments, investigated with mutants engineered for stability, Biochemistry. 37 (38) (1998) 13120–13127.
and H-bond interactions stabilize antigen-antibody interfaces, Proteins. 82 (9) (2014) 1734–1746.
[79] H.J. Hsu, K.H. Lee, J.W. Jian, H.J. Chang, C.M. Yu, Y.C. Lee, et al., Antibody variable domain interface and framework sequence requirements for stability and function by high-throughput experiments, Structure. 22 (1) (2014) 22–34.
[80] C. Tu, V. Terraube, A.S. Tam, W. Stochaj, B.J. Fennell, L. Lin, et al., A combination of structural and empirical analyses delineates the key contacts mediating stability and affinity increases in an optimized biotherapeutic single-chain Fv (scFv), J Biol Chem 291 (3) (2016) 1267–1276.
[81] S. Warszawski, A. Borenstein Katz, R. Lipsh, L. Khmelnitsky, G. Ben Nissan, G. Javitt, et al., Optimizing antibody affinity and stability by the automated design of the variable light-heavy chain interfaces, PLoS Comput Biol 15 (8) (2019), e1007207.
[82] G.A. Dalkas, F. Teheux, J.M. Kwasigroch, M. Rooman, Cation-pi, amino-pi, pi-pi, [83] B.J. Klein, K.R. Vann, F.H. Andrews, W.W. Wang, J. Zhang, Y. Zhang, et al., Structural insights into the pi-pi-pi stacking mechanism and DNA-binding activity of the YEATS domain, Nat Commun 9 (1) (2018) 4574.
[84] J.N. Haidar, Q.A. Yuan, L. Zeng, M. Snavely, X. Luna, H. Zhang, et al., A universal combinatorial design of antibody framework to graft distinct CDR sequences: a bioinformatics approach, Proteins. 80 (3) (2012) 896–912.
[85] K. Decanniere, A. Desmyter, M. Lauwereys, M.A. Ghahroudi, S. Muyldermans, L. Wyns, A single-domain antibody fragment in complex with RNase a: non-canonical loop structures and nanomolar affinity using two CDR loops, Structure. 7 (4) (1999) 361–370.
[86] B. Al-Lazikani, A.M. Lesk, C. Chothia, Standard conformations for the canonical structures of immunoglobulins, J Mol Biol 273 (4) (1997) 927–948.
[87] N. Wang, W.F. Smith, B.R. Miller, D. Aivazian, A.A. Lugovskoy, M.E. Reff, et al., Conserved amino acid networks involved in antibody variable domain interactions, Proteins. 76 (1) (2009) 99–114.
[88] C.A. Janeway Jr., T.P. M. Walport, et al., The structure of a typical antibody molecule, in: Immunobiology: The Immune System in Health and Disease, 5th edition, Garland Science, New York, 2001.
[89] A. Vangone, A.M. Bonvin, Contacts-based prediction of binding affinity in protein-protein complexes, Elife. 4 (2015), e07454.