

790397 Protein Engineering

6 Questions, 5 points each

27-30 points: sehr gut (1)

22-26 points: gut (2)

18-21 points: befriedigend (3)

15-17 points: genügend (4)

Don't forget to write your name and matriculation number on every page

Exam questions (Nr. of occurrence)

Chapter 1 Antibodies (Rüker)

Describe how to design a linker between two protein domains. Using an scFv as example, how can the necessary length of the linker be determined? What happens if the linker is very short? Which types of amino acids are well suited in such linkers, and which are less well suited? (1)

Describe and compare mouse antibodies, chimeric antibodies, humanized antibodies and fully human antibodies. Which methods are available today for the generation of fully human antibodies? (2)

Chapter 2 BsAbs and ADSC (Gordana)

Use of bispecific, biological use ? What is the purpose? How can you utilize multispecific binding of an antibody in a biological aspect ?

Chapter 3 Alternative binding scaffolds

Describe the molecular properties of alternative antigen-recognition scaffolds (other than antibodies). Which methods can be applied in engineering a protein into an antigen recognition scaffold? (3)

Which are common properties of alternative antigen binding scaffolds? What are the main differences to antibodies? Briefly describe two examples of alternative binding scaffolds. (3)

Chapter 4 Enzyme Evolution (Peterbauer)

The accumulation of mutations causes destabilization of a protein and/or compromises solubility and foldability. What strategies do we have to counter that? (2)

Chapter 9 Covid-19 vaccines (Rüker)

Describe and compare the strategies that are currently applied for the generation of vaccines against SARS-CoV-2. (2)

Chapter 10 High throughput screening

Describe the major characteristics and prerequisites of "screening by selection" (1)

What are the major characteristics and limitations of screening assays performed directly on agar plates? (2)

Chapter 11 Fusion proteins

Describe the main fields of applications of fusion proteins in research and therapy including 2 examples from each field. Which parameters are important for the design of a fusion protein?

Fusion proteins are commonly used for many different purposes. Briefly describe fusion proteins designed for (1) therapeutic purposes, (2) facilitation of protein purification and/or detection. Please refer to two examples in each category. (2)

Chapter Peterbauer

Which molecular strategies can be applied to fight allergy and which kinds of proteins are used for this purpose? Name a few of them. Describe the mechanism of action of one of the antibodies that are relevant in the field.

Chapter 1: Antibodies

Introduction to antibodies

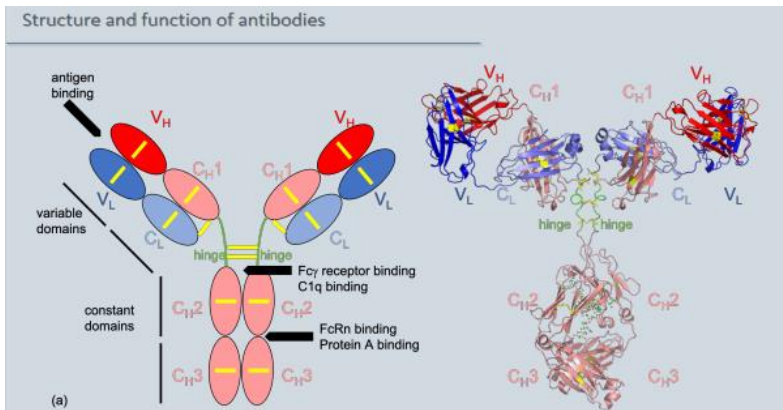


Figure 1 Structure of antibody

Antibody consists of a 2 heavy chains (pinkish, red) and 2 light chains (blue). Heavy chain consists of 4 domains (V_H , C_H1 , C_H2 , C_H3) and light chains of 2 domains (V_L , C_L). Yellow shows cysteine bridges. N-terminal end from both chains are where the antigen binds. Hinge (Gelenk) most flexible part of antibody.

Antibodies can bind below the hinge region to $Fc\gamma$ receptor and to C1q (first component of complement system). And at the interface of $CH3$ and $CH2$ $FcRn$ (neonatal Fc receptor, which is responsible for the long in vivo half life of antibody). binding site can be found. Protein A binding site can also be found there, which is relevant for the purification of antibodies

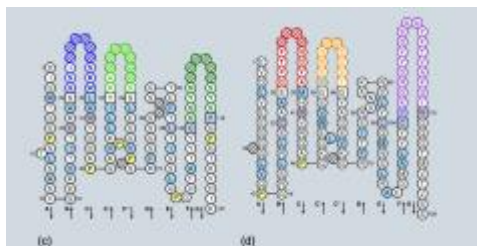
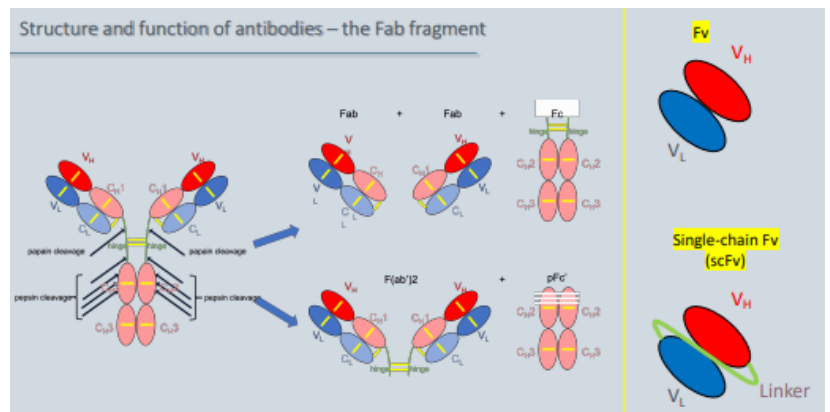


Figure 2 CD3 regions, antigen binding site

Fig 2 above shows coloured the CD3 regions of heavy and light chain which are actually binding to the antigen.

Figure 3 Fragmentation of an antibody

For engineering reasons nomenclature of different fragments is important. Scientists were able to split antibody into different fragment, on which they can work on.



What can be engineered?

- **Binding affinity and specificity**

Look at interface residues between antibody and antigen

Decide upon rationally designed mutations or in vitro directed evolution using library strategies

- **Valency**

Full IgG, Fab, $F(ab')_2$, scFv, diabody, other formats...

- **Effector functions**

Engineer contacts between Fc and $Fc\gamma$ -receptor

- **In vivo half-life**

- Engineer contacts between Fc and $FcRn$ -receptor

Q: Describe how to design a linker between two protein domains. Using a scFv as example, how can the necessary length of the linker be determined? What happens if the linker is very short? Which types of amino acids are well suited in such linkers, and which are less well suited?

A breakthrough was: it was finally possible to produce Fv regions recombinantly. Due to the fact that these two domains were not stable enough, they engineered a linker in between the domains. (scFv single chain Fv)

The length of a linker was determined by thinking of the Fv region as a sphere. And when you want to make a linker between the C-terminal of the VH with the N-terminal of VL, you have to “go” around the complete sphere. First determine the circumference of sphere $2\pi r$. Take the half of it and divide by the length of an amino acid.

So they found out, that the linker needs a length of about 15 AA

Typically small, neutral, characterless, flexible amino acids are used for linkage:

mainly glycine (no sidechain-only H atom, really small and flexible), and some serine to increase hydrophilicity and better solubility (also aids stability)

(**NO GOs**: cysteine, because of S, when

it doesn't make disulphide bonds within the molecule it tends to make with others which

lead to aggregation. Also hydrophobic AAs are No Gos because they even decrease solubility. Interesting fact: G, S linkers are also found in nature (protein 3 of phages S and G AAs link two subunits together)

If it's too short: you get fragments which can bind to more than one antigen. (bivalent Diabody, trivalent Triabody), because VH and VL of different Fv molecules will “bind” to one another.

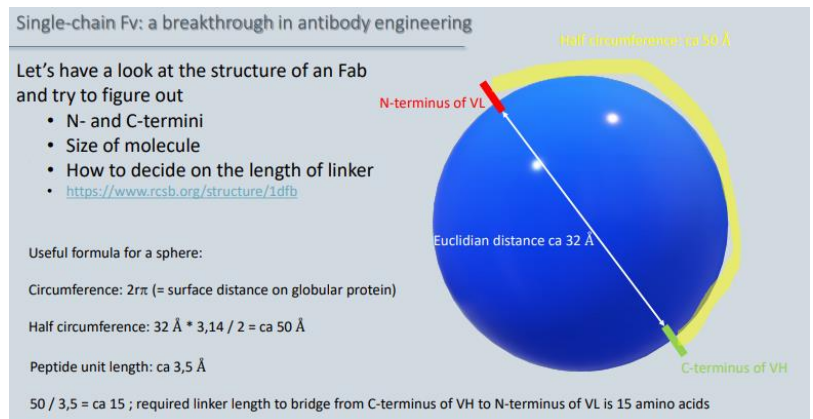


Figure 4 Determine length of linker

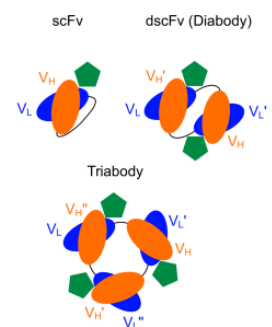


Figure 5: What happens if linker is short

Q: Describe and compare mouse antibodies, chimeric antibodies, humanized antibodies and fully human antibodies. Which methods are available today for the generation of fully human antibodies?

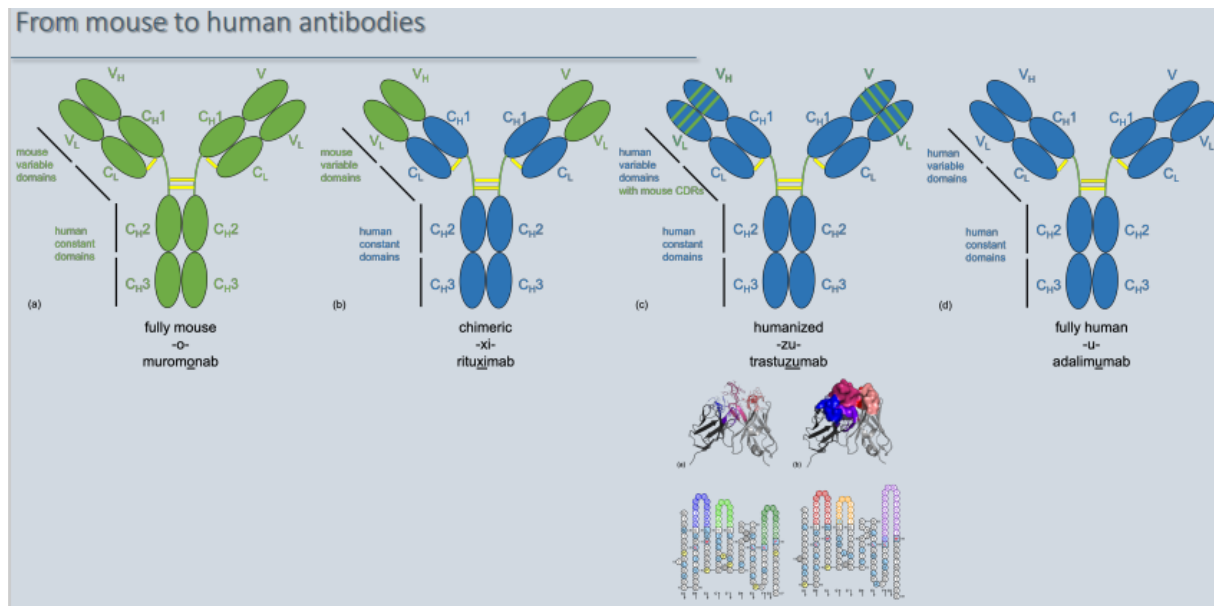


Figure 6 Evolution of mouse to human antibody

In the beginning antibodies were made in mouse. They immunized mice with antigen and isolated antibodies from it. The problem was that if you inject a mouse antibody into patient you will get a strong immune response. So chimeric antibodies were engineered to make them more human like. Chimeric antibodies are hybrids; VH from mouse is fused to human CH1, CH2, CH3 part of human and VL fused to human CL. ("xi" do chimeric antibodies have in their names). Further step to get even more humanlike antibodies is to take only CDR regions from the mouse and put them on human antibody. These are called humanized antibody and named with "zu"

2 ways to get fully human antibodies:

1. Transgenic mice are used, in which DNA is manipulated to produce fully human antibodies. These mice can be immunized by any antigen and it will express fully human antibodies.
2. Or use huge human antibody libraries containing billions of human antibody sequences displayed and phage or yeasts.

Chapter 2 BsAbs (BispecificAntibodies) ADCS (AntibodyDrugConjugates)

Q: Use of bispecific, biological use? What is the purpose? How can you utilize multispecific binding of an antibody in a biological aspect?

Bispecific antibodies combine specificities of two epitopes or two antigens in one molecule. One agent that can do 2 things at the same time.

combinatorial bsAbs: combine two functions. if you need to block multiple ligands where you normally apply 2 or more antibodies; you can use one single bsAbs

obligate bsAbs: effects can only be achieved by the physical linkage of the two specificities in a single molecule and not by their mixture

Effects can be spatial (räumlich) brings together two different celltypes together (or enzyme with substrate) or effects can be temporal: where the binding of the two domains proceeds in sequential manner and antigen interact of one arm depend on binding on the other one, either due to kinetic or steric reasons

Examples

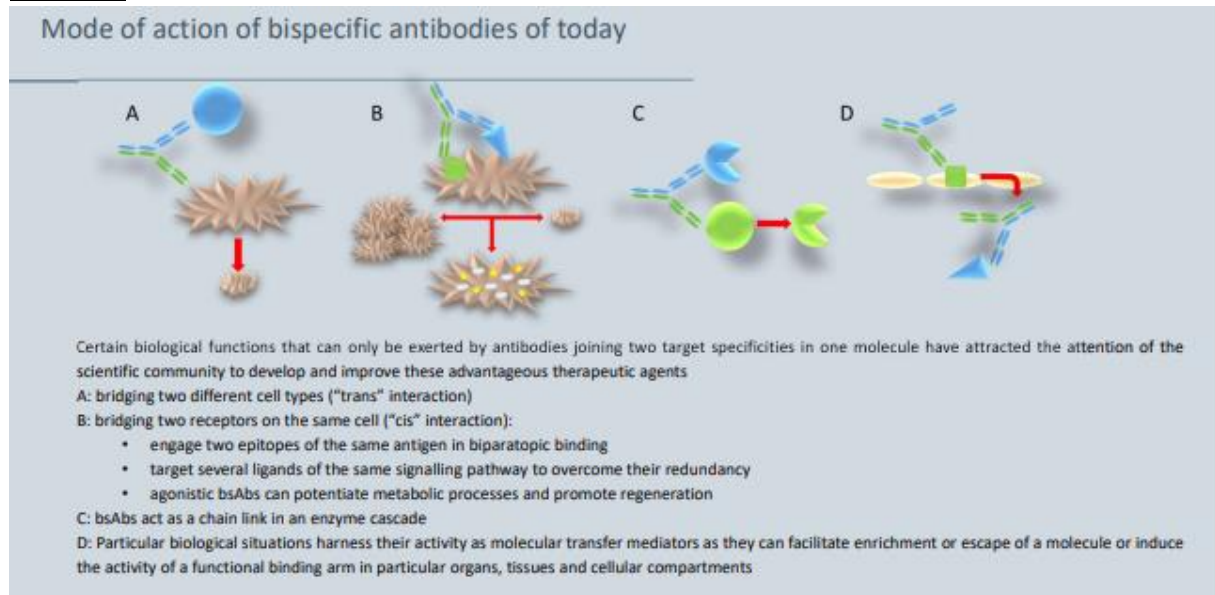


Figure 7 bispecific Antibodies

Type A: examples for trans interaction

Catumaxomab: First approved bispecific antibody, consists of mouse IgG2a and rat IgG2b

First Target: Ep-CAM (famous tumor receptor)

Second Target CD3 from T-cells

and actually third target by binding CH to Macrophages etc via Fc receptor

BiTEs (Bispecific T-cell engagers)

Blinatumomab

for treatment of B-cell Lymphoma, serial killers

consists of two singlechain variable fragments where one binds to CD3 (Tcell) and the other binds to CD19 (tumor cell) with a flexible linker between them.

When antigens are bound, T cell activation is triggered and releases cytotoxic granules and cytokines and Tcell proliferation and finally lysis of tumor (B-cell)

Emicizumab

Patients with haemophilia tend to bleed more than healthy patients. This is due the fact of lacking Factor VIII. Emicizumab mimicks Factor VIIIa(activated form of Factor VIII) and binds simultaneously to Factor IXa and Factor X. Patients treated with Emicizumab bleed less.

Molecular Trojan horses: Tackling the blood-brain barrier (BBB)

ADCs

Monoclonal antibody and drug are linked by a linker

Antibody binds to the binding site on tumor cells and is internalized. As long as drug is linked it stays inactive. Drugs on ADC are often 100-1000x more toxic than conventional chemotherapeutics. When ADC is in the cell and then degraded, drug is released and cause cell death.

Types of linker

- non-cleavable

increased plasma stability, larger therapeutic window, reduced off-target toxicity

- cleavable linker

acid-cleavable linker, remain stable at neutral pH but undergo hydrolysis and release the cytotoxic drug in acidic environment of the cellular compartments

- reducible linker

take advantage of the difference in reduction potential in the internal compartment versus plasma

- enzyme cleavable linker

take advantage of the abundance of hydrolytic enzymes with the specificity to recognize the sequences of peptides or patterns of carbohydrate.

Linker attachment sites

Lysin, Cystein, Glutamine (enzyme-mediated, using transglutaminase)

Antibody drugs as weapon against Covid-19

Donald Trump got treated with REGN-COV2 (a cocktail of two monoclonal antibodies)

Chapter 3: Alternative binding scaffolds

Q: Which are common properties of alternative antigen binding scaffolds? What are the main differences to antibodies? Briefly describe two examples of two examples of alternative binding scaffolds? Advantages and short comings

Everything that has been invented besides antibodies to allow antigen binding. These are very different proteins but with common concept in their engineering.

Properties of suitable alternative scaffold protein

- stable folding unit with sufficient thermostability (over 60°C)
- well expressed/displayed
- diversified residues intended for antigen binding form a confluent binding surface
- they are typically smaller than normal antibodies (good tumor penetration, good expressibility)

Display methods - directed evolution of antigen binding proteins

Typically, first a great library of possible binding scaffolds is created, and one is selected (<-directed evolution)

Rational design is rarely used.

Phage display: different possible binding scaffolds as fusion to coat protein of phages

advantage: good for simple peptides, high stability of phage particles, straightforward procedure

disadvantage: limited by transformation efficiency of E.coli, "only" 10^7 - 10^9 clones, due to prokaryotic nature translation and translocation machinery, not all protein fragments can be displayed (not glycosylated etc.)

Yeast display: displayed on surface of *s. cerevisiae*

advantage: better expression of mammalian proteins

disadvantage: even "smaller" libraries 10^6 - 10^8

mRNA ribosome display: no living cell needed, but mRNA not stable molecule you need professional laboratories.

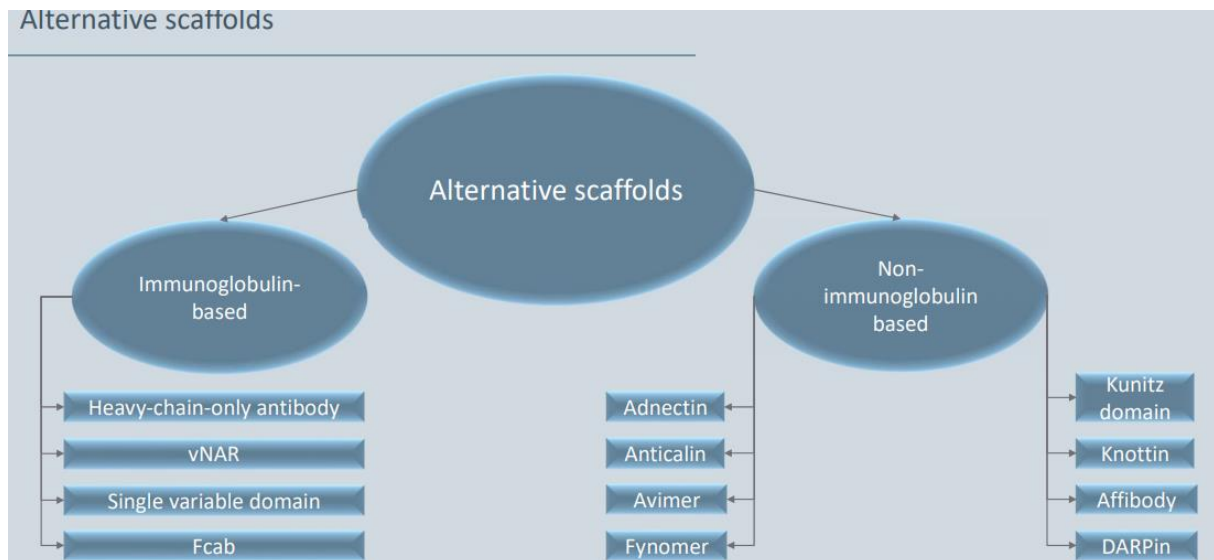


Figure 8 Types of alternative scaffolds

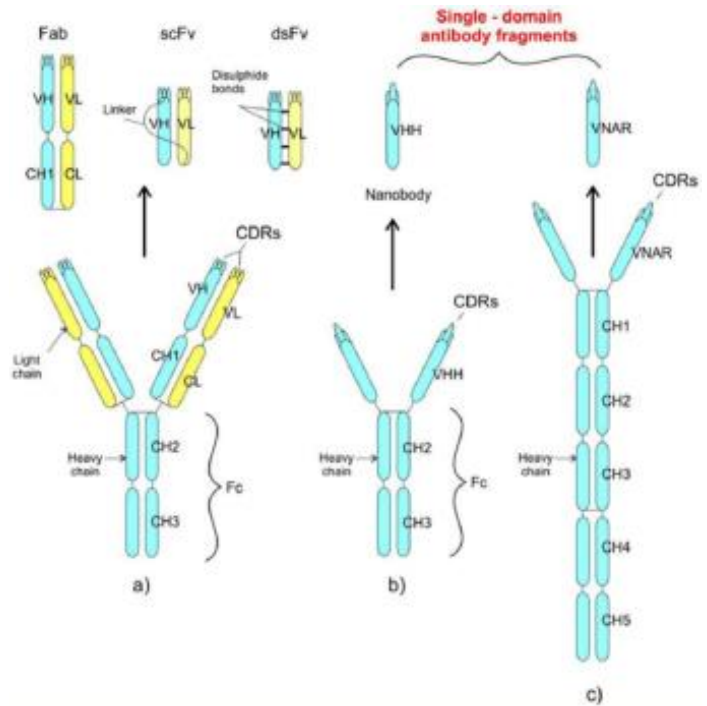
Immunoglobulin-based

Figure 9 Immuno globuline based alternative scaffolds

Heavy-chain-only antibody

1995 Scientists accidentally found in camelids Heavy-chain-only antibodies. This was the first time paratope is formed by only a single entity. How is it possible that one entity can have such a good affinity? Because of its elongated CDR3 loops. Due to their small size and elongated CDR3 loop they can address cryptic epitopes such as G-Protein coupled receptors, which are difficult to address using monoclonal antibodies “cleft binding”

Due to high sequence homology between VH and VHH domains, humanization of camelid VHH domains is a feasible method for the generation of variants which are potentially less immunogenic than the wildtype domain. They were also able to demonstrate that these humanized variants are amenable to loop-grafting of other CDR structures, yielding a general platform for humanization



vNAR (variable novel antigen receptor)

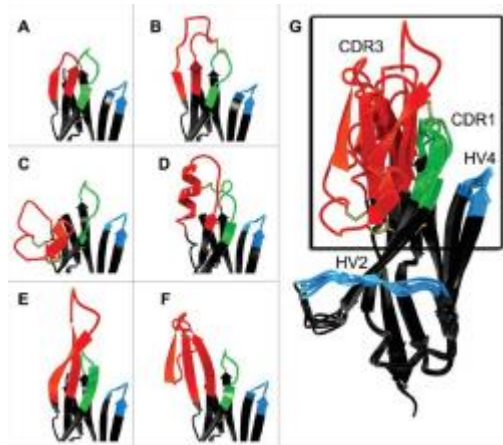


Figure 10 vNAR

found in sharks, different in the number of constant domains. Compose of 5 constant domains and a single variable domain which is connected via hinge-like linker. The variable domains of IgNAR antibodies are termed vNAR domains. They exhibit improved solubility due to the lack of light chain (VL) interaction partner. The binding interface, which would normally mediate VL chain pairing, therefore comprises an increased amount hydrophilic AA residues at positions that usually exhibit hydrophobic moieties.

They are extraordinary thermostable and if you heat denature them they simply fold back into native conformation afterwards. When taking a closer look at their structure it is noticeable that they are lacking CDR2 loop. Instead they comprise unusually long CDR3 binding sites which, for the most part, are responsible for antigen binding. The much shorter CDR1 loop is likewise involved in mediating antibody-antigen interactions, although in lesser extent. Besides CDR1 and CDR3, vNAR domains comprise two hypervariable loops termed HV2 and HV4

Single variable domain

only VH single domains represent a promising class of fragment being used as therapeutic modalities very small domain → very good for tumor penetration.

VH tend to aggregate, engineering of VH improve its properties, has a potential of being used of antigen binding scaffold

fully human vh domain can be constructed that are not only stable and well expressed but also rival the cleft binding properties of camelid antibodies

Fcab

Fc fragment of antibody is mutagenized to gain antigen binding properties.

advantage: very low immunogenic reactions.

Non-immunoglobulin based alternative scaffolds

Categories based on the location of amino acids that mediate ligand binding

- I) those with the ligand-binding amino acids in exposed loops (A-F) and
- II) those with these amino acids scattered in secondary structural motifs, such as α -helices (G-H)

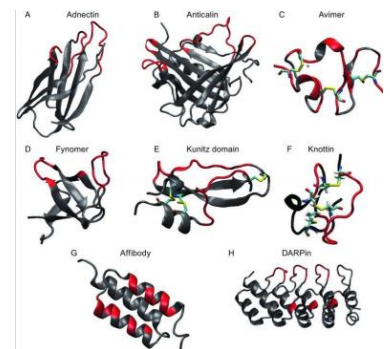


Figure 11 Alternative scaffold categories

Adnectins

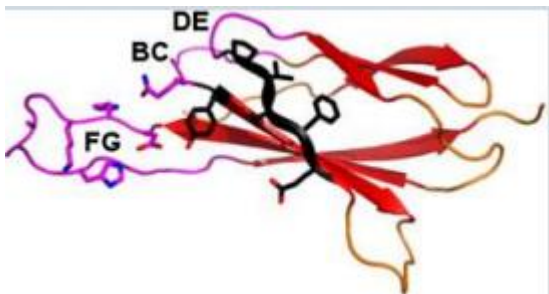


Figure 12 Structure of Adnectins

ligand binding AA are found in loops. Is based from the tenth domain of fibronectin type III (a human extracellular matrix protein). β -sandwich fold Quite similar to the IG fold (7 β - sheets, 6 loops which can be modified. FG-loop can even stand large insertions.

Anticalin (from butterfly)

Anticalin is a protein fragment derived from lipocalins, a class of secreted proteins that typically transport hydrophobic compounds. Adopts a conserved β -barrel structure consisting of eight anti parallel β -strands wound around a central axis and contains 160-180 AA (β -barrel).

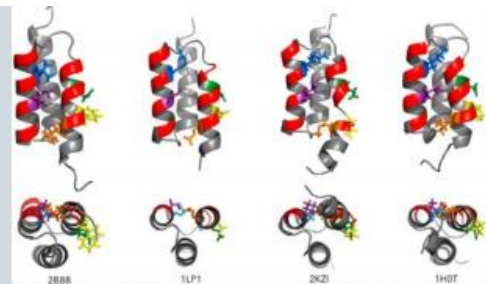
Anticalins are not glycosylated and do not possess any disulfid bonds, easily expressed in yeast and e.coli.

Half-life can be extended by site directed PEGylation (increases hydrodynamic volume of molecule).

Is already used to bind to VEGF-A.

Affibodies

- Affibodies are protein fragments derived from the Z-domain of the Ig-binding region of *Staphylococcus aureus* protein A which adopt a three-helix bundle motif and contain no cysteines
- These molecules possess high thermal and proteolytic stability and can be easily expressed in *E. coli*
- The ligand-binding surface is composed of 13 solvent-accessible residues scattered among two of the helices
- The small size (58 amino acids, 7 kDa) of affibodies allow them to be produced by chemical synthesis
- Affibodies exhibit rapid extravasation and rapid tumor penetration and unbound affibodies are quickly cleared from healthy organs and tissues, making them promising reagents for radionuclide imaging
- The affibody ABY-025 was engineered via phage display and affinity maturation to bind HER2 with low picomolar affinity and used for PET-imaging



Affibody structure with constrained sites highlighted. Side chains are shown for Q9 (orange), N11 (yellow), E15 (green), R27 (blue), and I31 (violet) in four determined affibody structures. 63,67,75,95 The other 10 classically diversified sites are colored red (Woldring, 2017)

Concerns of alternative scaffolds

Immunogenicity

- all non-host proteins are potentially immunogenic and carry the risk of being rejected by the host
- each individual case needs to be evaluated independently
- molecules that are potentially immunogenic are being largely developed for short-term imaging rather than for therapeutic applications

Short *in vivo* half-life

- most protein fragment therapeutics fall below the 70 kDa threshold for glomerular filtration
- several strategies have been developed to extend the protein fragment serum half-life, such as PEGylation and association or covalent conjugation with serum albumin or an antibody Fc domain
- none of these strategies are able to extend the half-life of protein fragments to that of native serum proteins such as antibodies and serum albumin, both of which have a serum half-life of ~21 days
- while a short half-life is not necessarily a disadvantage for the treatment of acute conditions, it represents a challenge for chronic or recurring illnesses.

Administration

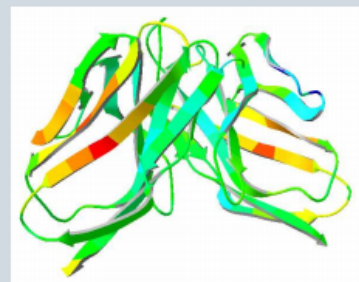
- Like antibodies, most protein fragments cannot be administered orally
- The acidic environment of the stomach, in addition to the activity of proteases in the stomach and small intestine, make it difficult for most protein fragments to make it through the digestive tract intact
- Exceptions are molecules that possess extremely a high proteolytic and chemical stability, such as knottins and potentially DARPins

Advantages

- the challenges associated with therapeutic mAb development are:
 - issues with host selection for generation of mAbs
 - humanization
 - high cost of manufacture
 - potentially poor tissue penetration
 - **Antibody-dependent enhancement** of viral infection
- mAb therapeutics have a high manufacturing cost and is currently mostly targeted to patients living in developed countries. The manufacturing cost of non-antibody protein fragments can be significantly lower, largely stemming from their production in microbial hosts. The development of non-antibody protein therapeutics may therefore be more economically feasible for patients, especially those suffering from viral or bacterial infection, in developing countries.

Immunogenicity

- Engineered recombinant proteins contain **non-natural** sequence contigs, which can be recognized as foreign by the patient's immune system
- **Every protein, either natural or not, is potentially immunogenic**, the immunogenicity does not depend on the sequence only
- There is **no reliable** prognosis of immunogenicity



Antibodies produced by the patient against a therapeutic protein (Anti Drug Antibodies, ADAs) can:

- Neutralize the therapeutic effect in the subsequent treatments
- Cause the formation of antibodies against autologous („own“) proteins, which can have fatal consequences
- Lead to anaphylactic reaction after readministering of the drug

How to engineer a protein into a scaffold?

by inserting sequences into loops or grafting onto other secondary structure (i.e. α -helices of (affibodies))

CD81 scaffold design:

Check Expressibility ✓ Improve expressibility

Check displayability ✓

Check stability ✓ Improve stability

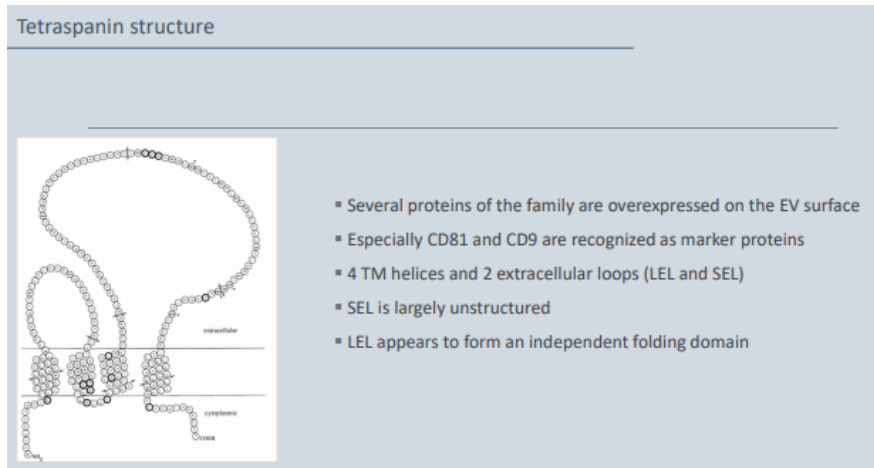
Create library with diversified antigenbinding sequences✓

Choose winner ✓

Introduction

Extracellular vesicles - communicator between the cells

- regulate information exchange between cells, especially with miRNAs
- biotechnologically interesting to modulate cell behaviour
- novel strategy as vector for drugs (effective transport of exogenously loaded compound)
- small size makes them better tolerated by host (immunotolerance)
- their membrane is enriched in tetraspanins (marker protein)
- the task: how to derivatize the EVs for directional targeting
- state of the art: overexpressed marker proteins have been addressed as acceptors for peptide grafts that enable specific recognition.



LEL consists of 5 helices and 2 disulfide bridges (also conserved). Consists of one conserved subdomain A,B,E , one variable subdomain "helices" C,D - are quite different in the family members of the tetraspanins.

CD81_LEL is suitable for grafting of shorter peptides, they started with one which recognizes transferrin receptor, so they make models how the molecule will look like if they insert the expanded the LEL. CD82 was modelled via software and they choose the one which looked most similar to CD81. They fit this model to CD81 to assure expressibility. Final model consisted of constant part from CD81 and variable part von CD82. They expressed in HEK and CHO to check expressibility.

Now stability needs to be checked by scanning with calorimetry. The higher the better. Was too low. But before they improve the stability they wanted to find out if structure can be correctly displayed on yeasts; The Sequence of CD81 LEL was cloned into pYD1 yeast display vector as a C-terminal fusion to AGA2p. Test for correct folding with structurally dependent anti-CD81 antibody. To check if it's in a correct conformation when its displayed on yeasts. And yes the antibody recognizes the structure. Display was possible, now they can stabilize structure.

First they diversify the scaffold, before they functionalize it for antigen binding By directed evolution approach with randomization of the wild-type scaffold and library selection for antigen-binding

molecules. To improve stability disulphide bonds were added to models. Then displayed. The ones which could still be displayed (check displayability) were purified and calorimetry measured and almost all led to higher stability.

- 1) when SEC one peak = perfect monomer
- 2) calorimetry shows high T_m (melting temp)
- 3) binding to the antibody good
- 4) CD circular dichroism same with wildtype

If all points good then make a library with binding sequence and check if protein is still stable after insertion. Choose scaffold winner. They choose 2 ; one with best stability one which was able to refold.

Now they inserted well known sequences from known antigens. Screen for antigen binding. Now mammalian cell display. Improve the clone. At last they tested, if cells which secrete laminin introduced to their clone via vesicle. They could show that vesicles were uptaken by these cells. Success.

Chapter 4 and 5 Enzyme evolution

Q: The accumulation of mutations causes destabilization of a protein and/or compromises solubility and foldability. What strategies do we have to counter that?

To counter that **gene duplication** or **upregulations** can be done. Also chaperons can help.

Gene duplication

- vehicle of functional divergence
- buffer for deleterious effects of mutations
- duplication increases protein dosis - reduces the stringency of purifying selection on individual proteins
- mildly deleterious mutations can accumulate to higher levels - accumulation of activity reducing mutations (ARMs) - compensated by duplications
- duplication in a different genomic location: often different regulatory regime
- new functions can arise from one protein being expressed in a different context
- duplication is often transient: relaxation of selection pressure or adaptive mutation that increases specific activity lead to loss of duplicate (pressure on genome streamlining/selective advantage of small genome)

Upregulation

Compensatory mutations leading to strong upregulations allows mutations to be fixed

Overexpression is costly and also often has deleterious side-effects.

Gene duplication is a better approach to compensate mutations

Chaperons

use of chaperons that lower the threshold for stability. overexpression of chaperons

Start with a more stable starting point

you either increase the stability afterwards or better: you start with a more stable starting point

evolution for increased stability first

you start first with a short run of directed evolution

Increase stability with stabilizing/compensatory mutations

Chapter 9 Library (Antibody and enzyme)

Principle: “genetic package” refers to physical coupling of information and function

Function: binding property (in case of antibodylibrary)

Information: necessary to replicate and produce desired selected binding protein

Principle of masterthief: We create a great library with possible binders to afterward find the right/best binder

A library contains many clones with different antibodies (e.g. scFvs). The challenge is to isolate the rare clones with the desired properties, e.g. binding to a specific antigen or epitope.

Important parameters in the library

Number of independent members:

- how many clones after transformation
- influenced by PCR bias, recombination bias, etc

Correctness how many correct clones in the library: unexpected stop, point mutations..

Productivity how many clones can produce a whole-length protein

Coverage:

- how does the actual size relate to theoretical size
- how many aminoacids are mutated
- relation of library size (nr. of independent members) to the theoretical size (theoretical diversity)

Sampling

- how many independent members are actually participating in the selection
- how many library entities are involved in selection
- how does that correspond to actual library size

Chapter 9 Covid-19

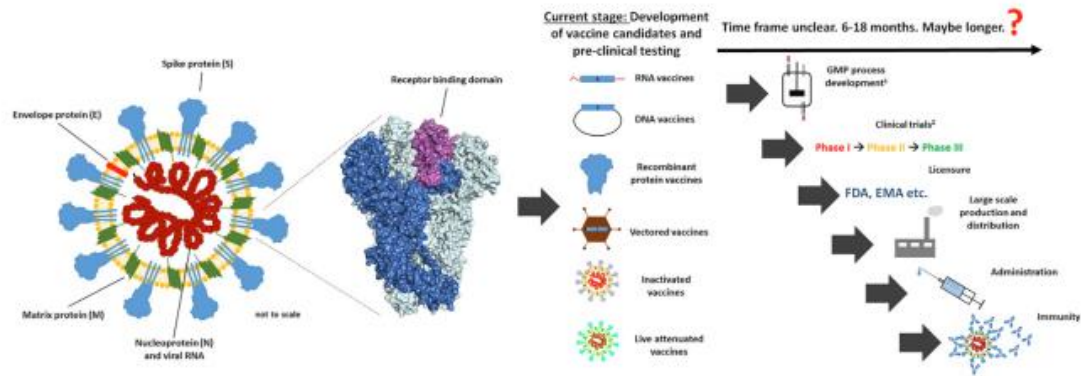


Figure 1. Overview of Potential SARS-CoV-2 Vaccine Platforms

The structure of a coronavirus particle is depicted on the left, with the different viral proteins indicated. The S protein is the major target for vaccine development. The spike structure shown is based on the trimeric SARS-CoV-1 spike (PDB: 5XL3). One trimer is shown in dark blue, and the receptor binding domain, a main target of neutralizing antibodies, is highlighted in purple. The other two trimers are shown in light blue. SARS-CoV-2 vaccine candidates based on different vaccine platforms have been developed, and for some of them, pre-clinical experiments have been initiated. For one mRNA-based candidate, a clinical trial recently started to enroll volunteers shortly (ClinicalTrials.gov: NCT04283461). However, many additional steps are needed before these vaccines can be used in the population, and this process might take months, if not years.¹ For some candidates, cGMP processes have already been established.² Clinical trial design might be altered to move vaccines through clinical testing quicker.

Platform	Target	Existing, Licensed Human Vaccines Using the Same Platform	Advantages	Disadvantages
RNA vaccines	S protein	No	No infectious virus needs to be handled, vaccines are typically immunogenic, rapid production possible.	Safety issues with reactogenicity have been reported.
DNA vaccines	S protein	No	No infectious virus needs to be handled, easy scale up, low production costs, high heat stability, tested in humans for SARS-CoV-1, rapid production possible.	Vaccine needs specific delivery devices to reach good immunogenicity.
Recombinant protein vaccines	S protein	Yes for baculovirus (influenza, HPV) and yeast expression (HBV, HPV)	No infectious virus needs to be handled, adjuvants can be used to increase immunogenicity.	Global production capacity might be limited. Antigen and/or epitope integrity needs to be confirmed. Yields need to be high enough.
Vectored vaccines	S protein	Yes for VSV (Ervebo), but not for other viral vectored vaccines	No infectious virus needs to be handled, excellent preclinical and clinical data for many emerging viruses, including MERS-CoV.	Vector immunity might negatively affect vaccine effectiveness (depending on the vector chosen).
Live attenuated vaccines	Whole virion	Yes	Straightforward process used for several licensed human vaccines, existing infrastructure can be used.	Creating infectious clones for attenuated coronavirus vaccine seeds takes time because of large genome size. Safety testing will need to be extensive.
Inactivated vaccines	Whole virion	Yes	Straightforward process used for several licensed human vaccines, existing infrastructure can be used, has been tested in humans for SARS-CoV-1, adjuvants can be used to increase immunogenicity.	Large amounts of infectious virus need to be handled (could be mitigated by using an attenuated seed virus). Antigen and/or epitope integrity needs to be confirmed.

Chapter 10 High Throughput Screening, microfluids

Q: Describe the major characteristics and prerequisites of “screening by selection” (1)

Q: What are the major characteristics and limitations of screening assays performed directly on agar plates? (2)

Screening by selection

Relative large library sizes can be screened. Advantage is that only variants with active enzyme is able to survive. Disadvantage is that not every enzyme can be screened by that method; only enzymes which are vital for survival of the host strain. Often custom-made recipient strain required: deletion strains for enzymes essential in biosynthetic pathways

Enzyme give host a capability

- utilization of sole C-or N-Source
- antibiotic/toxin resistance
- pH/heavy metal/salt tolerance

Screening is done by colony diameter directly on plates - catalytic efficiency of engineered/evolved enzyme relates to host viability. Digital camera + image analysis software allows identification and isolation of positives. Throughput is limited by expression organism and transformation efficiency for library

22x22 cm plates - more than 12000 e coli colonies possible

several 100000 colonies can be screened with manageable effort and time

10^6 , 10^7 per round

Plate assay screening

Screening by clearing halos on plates, color reactions/fluorescence with substrate analogs in overlay or sprayed on. Limited to secretory enzymes or inclusion of permeabilization step. Very simple operation. Screening of smaller library sizes than by selection. Limited dynamic range; means further analytics have to be made.

Example: *S.cerevisiae*. yeast used as expression system

Enzyme of interest is expressed on turbid medium. if yeast has enzyme which can degrade, halos occur around colony

Problem: further screenings are needed. Halos does not relate to actual activity. Halos can also occur by enzymes with low activity but high in concentration

Chapter 11 Fusion Protein

Describe the main fields of applications of fusion proteins in research and therapy including 2 examples from each field. Which parameters are important for the design of a fusion protein?

Fusion proteins are commonly used for many different purposes. Briefly describe fusion proteins designed for (1) therapeutic purposes, (2) facilitation of protein purification and/or detection. Please refer to two examples in each category.

A Fusion Protein that consists at least of two domains that are encoded by separate genes that have been joined so that they are transcribed and translated as a single unit, producing a single polypeptide. A protein fusion in which two different proteins are linked through peptide bonds; usually genetically engineered. Chimeric antibodies may have the Fab fragment from one species fused with the FC fragment from another.

The technique of creating fusion proteins has been extended to other fusion partners, and additional uses have been developed for the fusion partner. Three of the most important uses of fusion proteins are:

- as aids in the purification of cloned genes

- as reporters of expression level
- as histochemical tags to enable visualization of the location of proteins in a cell, tissue, or organism

Recombinant fusion proteins are created artificially by recombinant DNA technology for use in biological research or therapeutics

Chimeric or chimera usually designate hybrid proteins made of polypeptides having different functions or physicochemical patterns.

Chimeric mutant proteins occur naturally when a complex mutation, such as:

- a chromosomal translocation
- tandem duplication
- retrotransposition

creates a novel coding sequence containing parts of the coding sequences from two different genes. Naturally occurring fusion proteins are commonly found in cancer cells, where they may function as oncoproteins. The bcr-abl fusion protein is a well-known example of an oncogenic fusion protein and is considered to be the primary oncogenic driver of chronic myelogenous leukemia.

Purpose of engineering

Novel recombinant technologies have made it possible to improve fusion protein design for use in fields as diverse as biodetection, paper and food industries, and biopharmaceuticals. Recent improvements have involved the fusion of single peptides or protein fragments to regions of existing proteins, such as N and C termini, and are known to increase the following properties:

- **Catalytic efficiency:** Fusion of certain peptides allow for greater catalytic efficiency altering the tertiary and quaternary structure of the target protein
- **Solubility:** A common challenge in fusion protein design is the issue of insolubility of newly synthesized fusion proteins in the recombinant host, leading to an over-aggregation of the target protein in the cell. Molecular chaperones that are able to aid in protein folding, may be added, thereby better segregating hydrophobic and hydrophilic interactions in the solute to increase protein solubility
- **Thermostability:** singular peptides or protein fragments are typically added to reduce flexibility of either the N or C terminus of the target protein, which reinforces thermostability and stabilizes pH range.
- **Enzyme activity:** Fusion that involves the introduction of hydrogen bonds may be used to expand overall enzyme activity
- **Expression levels:** addition of fusion fragments such as maltose binding protein (MBP) or small ubiquitin-like molecule (SUMO), serve to enhance enzyme expression and secretion of the target protein
- **Immobilization:** PHA synthase, an enzyme that allows for the immobilization of proteins of interest, is an important fusion tag in industrial research (PHA has the habit to collect in granules)
- **Crystal quality:** Crystal quality can be improved by adding covalent links between proteins aiding in structure determination techniques

Examples

SUMO (small ubiquitin-related modifier)

- small globular (approx. 100 AA and 12 kDa)
- although very little sequence identity with ubiquitin (less than 20%), it has nearly identical structural fold.
- SUMO has a unique N-extension that other ubiquitin-like molecules do not have

Fusions result in many beneficial function, such as protein stability, nuclear-cytosolic transportation and transcriptional regulation

Typically, only a small fraction of a given protein is SUMOlyated and this modification is rapidly reversed by the action of deSUMOlyating enzymes. SUMOlyation of target proteins has been shown to cause a number of different outcomes including altered localization and binding partners:

Biotechnological use of SUMO tag

fusion of protein of interest with SUMO increase the solubility, because SUMO has an external hydrophilic surface and inner hydrophobic core, which may exert a detergent like effect on other insoluble proteins

SUMO also act like a nucleus of folding → SUMO- tagged proteins fold more properly

Another advantage of SUMO is that it can be easily cleaved

VHH expressed in E.coli: SUMO-tag lead to higher yield of VHHs in soluble form and did not affect the antigen-binding activity of the VHHs.

Affinity purification tags

Staphylococcus Protein A

HIS tag

SPY-tag

Tag that binds forever

If a protein of interest is bound to the spy tag it is bound forever. Streptococcus pyogenes fibronectin-binding protein FbaB contains a domain with a spontaneous isopeptide bond between Lys and Asp (covalent bond). By splitting this domain and rational engineering of the fragments, we obtain the SPY-tag which formed an amide (covalent) bond to its protein partner (spy catcher)

Asp on spytag and lys on spy catcher are bound covalently together.

Reporter tags

GFP is fused to the protein of interest to visualize it directly or quantitated using a luminometer. advantage no substrate is needed (luciferase another possible reporter fusion partner needs a substrate), simple in use, not expensive,

β-can structure with the chromophore buried inside

Fusion proteins in therapy

Fusion proteins are typically composed of the extracellular domains of native transmembrane proteins, such as cell surface receptors, linked to another molecule

In most cases the linker that has been used has been the Fc portion of human Ig, which achieves the desired effect of enhancing the pharmacokinetic properties of the construct. The Fc portion of the fusion receptor can be maintained to be functional and bind to complement and to cell surface Fc receptors. This allows effector function, and has been used for the compounds thanercept and alefacept. Alternatively, the construct can be engineered such that the Fc piece is unstable to execute effector functions, as in the case of abatacept, which does not activate complement

As their primary mechanism of action, fusion receptors competitively inhibit the binding of a ligand to its specific receptor and thereby prevent downstream effects, such as activation of inflammatory cascade. However, ligation of certain molecules by fusion, receptors may induce other functions as well.

Fusion of HSA (Human Serum Albumin) or FC portion led to half-life-extension

PEG (Polyethyleneglycol) attached to the drug is able to mask the drug from host's immunesystem. Can also provide water solubility to hydrophobic drugs and proteins

- improved halflife
- reduced immunogenicity
- increased solubility
- minimized aggregation

Summary: strategies of half-life extension

(Strohl, 2015)

- Genetic fusion of the pharmacologically active peptide or protein to a naturally long-half-life protein or protein domain (e.g., Fc fusion, transferrin [Tf] fusion, or albumin fusion)
- Genetic fusion of the pharmacologically active peptide or protein to an inert polypeptide, e.g., XTEN (also known as recombinant PEG or "rPEG"), a homo-amino acid polymer (HAP; HAPylation), a proline-alanine-serine polymer (PAS; PASylation), or an elastin-like peptide (ELP; ELPylation)
- Increasing the hydrodynamic radius by chemical conjugation of the pharmacologically active peptide or protein to repeat chemical moieties, e.g., to PEG (PEGylation) or hyaluronic acid
- Significantly increasing the negative charge of fusing the pharmacologically active peptide or protein by polysialylation; or, alternatively, (b) fusing a negatively charged, highly sialylated peptide (e.g., carboxy-terminal peptide [CTP; of chorionic gonadotropin (CG) β -chain]), known to extend the half-life of natural proteins such as human CG β -subunit, to the biological drug candidate
- Binding non-covalently, via attachment of a peptide or protein-binding domain to the bioactive protein, to normally long-half-life proteins such as HSA, human IgG, or possibly transferrin
- Chemical conjugation of peptides or small molecules to long-half-life proteins such as human IgGs, Fc moieties, or HSA.

Other therapeutically important protein fusions

- Fusions of antibodies with toxins (diphtheria toxin, pseudomonas toxin)

Conclusions

Three of the most important uses of fusion proteins are:

- as aids in the purification and of detection of recombinant proteins
- as aids for expression and solubility level of recombinant proteins
- as reporters of expression level
- as histochemical tags to enable visualization of the location of proteins in a cell, tissue, or organism

Therapeutic applications of fusion proteins include:

- Strategies for half-life extension
- Extending functionalities of therapeutic proteins: ADC, biological toxin, radioconjugates

GOOD LUCK!