

Linkers: a synergistic way for chimeric proteins

Dharti Patel¹, Dhanya Menon¹, and Darshan Patel¹

¹P D Patel Institute of Applied Sciences

October 6, 2020

Abstract

Linkers are short or long stretch of amino acid sequences used to create chimeric proteins. Linkers can adopt various structures in a biological environment which is a deciding factor for the functionality of a chimeric protein. It is essential to generate fully separated proteins in the chimeric protein and an effectual linker design increases the stability and activity of the protein. In this review, we summarize the types of linkers and strategies for designing linkers in chimeric protein generation. Appropriate linker design results in the creation of recombinant fusion protein which can be employed in biochemical analysis, drug targeting, pharmaceuticals, diagnostic, and industrial and biotechnological applications.

Introduction

An increase in crystallographic data of protein structure reveals the dynamism of protein multi-functionality. Many proteins show multiple or sequential functions where the connector region or linker that connects two domains of individual proteins helps to generate multi-functionality as shown in figure 1. An engineered biomolecule comprising fragments from more than two genes and a single multifunctional construct is produced with the help of a linker which is referred to as fusion protein (Zhao et al., 2008a). Thorough knowledge of such a connector or linker region is necessary, to produce an efficient chimeric protein. Such chimeric proteins are of immense use at the industrial level by reducing reaction time, improving protein properties, and faster product formation. They can also be used in medical applications such as cancer therapeutics, drug targeting and imaging.

Several reports have shown that the chimeric proteins have improved the efficacy of the proteins along with multi-functionality. It has been observed that linking BFP(Blue fluorescent protein) and GFP(Green fluorescent protein) has improved the efficacy of both the proteins by 26.5%(Ryoichi Arai et al., 2004). Similarly, G-CSF(Granulocyte-colony stimulating factor)and transcription factors combined with the help of the linker had shown improvement instability(Bai & Shen, 2006a). Industrially relevant enzymes, β -glucanase, and xylanase linked with a linker showed increased stability(Lu & Feng, 2008a). When a 15 amino acid long linker was used to fuse protein A with coat protein, this resulted in functional virions(Werner et al., 2006). G-CSF showed almost 10-fold higher bioactivity through the (A(EAAAK)4ALEA(EAAAK)4A) helical linker insertion(N. Amet et al., 2009). HSA (Human serum albumin) with (Interferon) INF- α 2b fused with the help of a linker showed enhanced half-life under in vivo conditions(Bai & Shen, 2006b).

The choice of a protein linker plays a crucial role in the successful construction of fusion protein. It is essential to keep two proteins or domains distinctly apart which do not allow much intra protein or domain interactions for chimeric protein generation. In many instances, a direct fusion of two genes results in an impaired activity of proteins(Tuna et al., 2019). Direct fusion often changes the ionic network of the proteins

which results in a non-functional protein. It also shows the capacity to enhance the soluble protein expression when expressed in a different expression host. An effective linker provides suitable space between two proteins which will decrease their intrusion, improve or restore folding or ultimately improve bioactivity(Chen et al., 2013b).

Several linkers have been identified and employed for chimeric protein production. In this review, we focus on various types of linkers and their selection criteria for the chimeric protein generation.

Fig.1. Linker strategy for chimeric protein construction

Criteria for selection of linker

In due course of evolution, living organisms have generated a variety of proteins within the biological system, displaying multi-functionality by the presence of multi-domains connected by a linker peptide. Biophysical analysis has revealed several types of naturally existing linkers with different sizes and shapes. These linkers exhibit secondary structure like α helix, β strand as well as coil and turn. α helix and coil/bend showed their highest presence in the multi-domain proteins while β strand and turn are in fewer numbers in the multi-domain proteins(Nick Pace & Martin Scholtz, 1998). In genetic fusion, Gly (Glycine), Ser (Serine), and Thr (Threonine) residues, which are found in the highest number in natural linkers, create efficient linkers(Reddy Chichili et al., 2013).

A suitable linker design involves definite interspace of interaction site at N- terminus or C- terminus where the fusion protein binds. Hydrophilicity and a certain degree of flexibility of linkers are essential for proper interaction in chimeric proteins. Hydrophilic linker provides benefit to the chimeric protein by enhanced solubility, free movement to each other, and resistance to proteolysis(Tollefsen et al., 2012). Amino acids with small side chains were found to be desirable linker constituents(Argos, 1990a). Glycine has a low preference for the formation of α -helix(Pace & Scholtz, 1998). The information on secondary structure and rigidity as well as the tendency of amino acid arrangement intended for natural linkers is helpful for empirical linker designing and it is customized for exacting applications.

The structure of the linker is dependent on the length and amino acid sequence of the linker. In general, 2 and 31 amino acid linkers are used for chimeric protein construction(Saadat, 2017). The length of the linker can be decided with the previous knowledge of the binding region. To enable efficient hydrolysis, specific inter-domain spacing is required because extremely short or long linkers would be interacting in catalysis(Ruiz et al., 2016a). Serine and threonine residues in linkers help to prevent linker collapse and protein flexibility and assist in conforming stabilization between flanking molecules. It is possible to implement a stretched conformation in the engineered chimeric protein (Crasto & Feng, 2000). Many proline and glycine residues exist in almost all linkers. Proline residues enable extended conformations by increasing rigidity(Receveur et al., 2002a). For appropriate orientation, additional flexibility between the two domains can be provided by glycine (Sammond et al., 2012). Influenza virus hemagglutinin peptide- GSGSGS, Hen egg lysozyme – GGGSLVPRGSGGGGS, myelin, moth Cytochrome c- GGGSLVPRGSGGGGS, and ovalbumin peptide-GSGSGS are some examples of linkers which are rich in glycine residues(van Rosmalen et al., 2017).

The length of linkers plays a significant role in enzyme action and thermal adaptation. Deleting or shortening the linker may have implications on enzyme activities(Ruiz et al., 2016b). Rates of stability, protein folding as well as domain-domain positioning are properties of the linker which is affected with variation in linkers length(Priyanka et al., 2013a). Entirely stretched positions in extended linkers tend to implement reduced conformations. Distance between the carbohydrate-binding module and the catalytic domain is essential for efficient catalysis. If a higher amount of hydrophobic residues are present it leads to the production of less stretched conformers(Receveur et al., 2002b). The highest turnover rates were achieved with more rigid variants. Though it is necessary to allow some flexibility in the linker's specific region, this indicates the importance of a balanced and proper arrangement of flexible and rigid portions in the linker for intact catalytic properties of enzymes.

Types of linkers

Structurally linkers can be divided into helical and non-helical (R A George & Heringa, 2003). They act as scaffolds to avoid uncomplimentary interaction among folding domains. Helical linkers are used as stiff spacers and extended structure helical linkers are used for separation of functional domains which greatly reduces their interference and allows fusion proteins to fold properly with enhanced stability (Chen et al., 2013b). Non-helical linkers are applied for structural stiffness as well as segregation of the linkers from the attached domains. Two domains are allied with a sequence of linkers for chimeric protein generation. Multifunctional fusion proteins holding domains from different proteins are proline-rich non-helical linkers (R A George & Heringa, 2003). The composition of the sequence for the linker has a considerable effect due to the folding strength of a chimeric protein (Robinson & Sauer, 1998a).

Physically linkers can be classified into three large groups; flexible linkers, rigid linkers, and in vivo cleavable linkers (Nagamune, 2017). Details of each type with examples are given in table 1. A diagrammatic representation of different types of linkers is represented in figure 2.

Figure 2. A diagrammatic representation of different types of linkers

Table 1. Summary of linker types

Flexible linkers

Generally, flexible linkers are used when an interaction or movement with a certain degree of the linked domains is demanded (Chen et al., 2013b). They can serve as passive linkers by retaining distance between functional domains. The length of the flexible linkers can be adjusted to allow for proper folding, improved stability, or to accomplish the optimal biological activity of the fusion proteins (Chen et al., 2013a). Flexible linkers are composed of small, non-polar (e.g. Gly) or polar (e.g. Ser or Thr) amino acids (Argos, 1990b) which permits mobility and offers flexibility for the connecting functional domains. Solubility, as well as stability improvement, can be achieved by linking proteins to other proteins with linkers that are rich in glycine residues and thus stable protein-protein complex are generated (Gadd et al., 2011) (Robinson & Sauer, 1998b). Proximity is increased by the preservation of natural interaction between two interacting proteins with the help of a linker (Priyanka et al., 2013b).

Properties of a protein, as well as their natural interaction, do not change when natural flexible linkers rich in glycine residues are used. Employment of Gly-rich linkers can be seen in imaging studies (Priyanka et al., 2013b). When proline residues are exchanged in place of glycine residues at the terminal, the degrees of independence are lost (Ruiz et al., 2016b). As a result, enzyme activity is decreased leading to lower catalytic efficiencies. Thus, flexible linkers provide a reduced catalytic activity to proteins (Ruiz et al., 2016a). Proteins having extra flexible linkers exhibit a reduced turnover rate due to lesser catalytic efficiency (k_{cat}/K_m) (Ruiz et al., 2016b).

The movement of the attached domains is essential for a definite degree of interaction. A mainly used common flexible linker is (Gly-Gly-Gly-Gly-Ser) $_n$ where “n” specifies the sum of repeats of the motif shown in figure 3(a) (Huston et al., 1988). These linkers are unstructured because of their flexibility; hence they provide partial domain separation. Less quantity of Gly, Ser, and Thr residues results in flexibility which permits mobility of the linking functional domains. For the reduction of linker protein interaction with protein function, serine polar residues are added (Evers et al., 2006).

Figure 3(a). Flexible linker showing a loop-like structure. (Gly-Gly-Gly-Gly-Ser) $_n$ present between V_H and V_L chain of the antibody.

Rigid linkers

Rigid linkers separate the functional domains more efficiently than flexible linkers (Gräwe et al., 2020). Loss of biological activity or poor expression yields is observed with flexible linkers. The length of the linkers can be adjusted by changing the copy number i.e. repetitive sequence to attain an optimal distance between

domains(Różycki et al., 2017). It is necessary to opt for rigid linkers when the spatial separation of the domains is necessary to preserve the stability or bioactivity of the fusion proteins. It is also observed that the turnover rate is higher in enzymes with rigid linkers(Ruiz et al., 2016a). Rigid linkers display rigid structures by implementing α -helical structures due to multiple Pro(Proline) residues (Chen et al., 2013a). Rigid linkers have been effectively useful for protein domains to retain a secure gap between them and to uphold their independent functions(Chen et al., 2013b).

Generally, domains from the same protein possess rigid linkers with Arg (Arginine), Pro, Gln (Glutamine), Thr, Glu (Glutamic acid), and Phe (Phenylalanine) amino acids. Structural studies on rigid spacers which are rich in proline residues are useful for the prevention of unfavorable interactions between the domains(Richard A. George & Heringa, 2002). An increase in rigidity of flexible linkers can be obtained by integrating probable N-linked glycosylation positions in it. An (EAAAK)_n A where (n=2-5) was the first experiential rigid linker designed by Arai et al. shown in **fig.3(b)**. Similar rigid linkers employed for the generation of chimeric proteins are (EAAAK)₃, A (EAAAK)₄LEA (EAAAK)₄A, (Alanine-Proline) (Ala-Pro)_n (10–34 AA). α -helical assemblies having many prolines are adopted by rigid linkers, thus exhibiting relatively stiff structures(R. Arai et al., 2001).

In a nutshell, rigid linkers are chosen in a situation when the stability or bioactivity of longitudinal parting is susceptible to reserve.

Φιγυρε.3(β) Ριγιδ λινκερ σηωωινγ αν α-ηελικαλ στρυςτυρε

In vivo cleavable linkers

In vivo, cleavable linkers are stable linkers. They covalently join functional domains together to act as one molecule throughout the *in vivo* processes(Dorywalska et al., 2016). This stable linkage between functional domains offers several benefits such as a prolonged plasma half-life (e.g. albumin or (Fragment crystallizable region) Fc-fusions). *In vivo* cleavable linkers are designed to be prone towards a particular protease by incorporating specific protease-sensitive sequences. Unlike the reduction of disulfide bond which happens rapidly in the blood circulation(Dai et al., 2010), the specificity of many proteases provides slower cleavage of the linker at controlled compartments(Poreba, 2020). *In vivo* cleavable linkers can be conjugated to antibodies by lysine and cysteine residues(Leung et al., 2020).

In vivo, cleavable linkers are beneficial in liberating the free functional domain *in vivo* (Chen et al., 2010a). Besides this, some additional benefits may be offered by linkers for fusion protein construction. For instance, biological activity enhancement, improvement in protein expression yield, and attaining required pharmacokinetic profiles(Tsuchikama & An, 2018). The disulfide linker designed by Chen et al(2013), (LEAGCK-NFFPR—SFTSCGSLE) is an example of an *in vivo*cleavable linker represented in **figure 3(c)** .

Figure 3(c). *In-vivo* cleavable linker

Role of linkers

The basic role of rigid and flexible linkers is to connect the functional domains. Chimeric proteins can accomplish the united function of merged proteins (Deane et al., 2004). Proteolytic resistance, stability, and solubility of chimeric proteins may be affected by the firmness of linkers(Costa et al., 2014). Due to this, an uncomplimentary interface inside aqueous solution among protein molecules and the linker is shortened. For the preservation of the biological activity of the fused complex, the length of the linker and amino acid composition should be significantly enhanced(Robinson & Sauer, 1998b).

Maximum conformational freedom is provided by poly-Gly linkers(Ruiz et al., 2016b). On contrary to it, a linker with (Alanine)_n,(Gly)_n,(Ser)_n where n=11,5,7 amino acids respectively in a haphazard area within linker provide extreme stability(Robinson & Sauer, 1998b). Gly rich flexible linkers are used for the advancement of protein folding with the epitope. A Gly rich flexible linker with 5 to 10 residues is used between a tagged protein and epitope to increase epitope sensitivity(Chakdar et al., 2016). Enhancing biological actions, improving production, plasma stability of tagged antibodies attaining a targeted or controlled drug

transfer are the advantages provided by linkers (Chen et al., 2013a). Efficient proteolytic cleavage in cells can be achieved by the accurate design of protein linker sequence (B. Wei et al., 2018).

The basic and significant function of recombinant chimeric proteins is to covalently link the functional domains through flexible linkers, rigid linkers, or cleavable linkers for releasing them under required conditions. Through protein engineering techniques, a flexible amino acid linker is used for linking two binding partners, to address unstable or unfolded proteins. Both intra-molecular, as well as inter-molecular protein-protein interactions, can be maintained through fusion proteins (Kingston et al., 2004).

The gap between two efficient domains is controlled through the length and structure of linkers, which is responsible for chimeric protein stability (McCormick et al., 2001). To construct dual-functional chimera of xylanase and β -glucanase, the α -helical linkers (EAAK)_n where (n=1-3) and flexible linkers have been tried and found to be useful (Chen et al., 2013a). Increasing linker length increased the thermostability of β -glucanase. As compared to flexible linkers, rigid linkers with an α -helical structure allow the protein to function and fold autonomously which is responsible to provide higher thermal stability (Lu & Feng, 2008b).

For recognition and binding of substrate, minimum elasticity is essential to avoid conformational heterogeneity through excessive structural disorder. In enzyme with a flexible polypeptide, the yield rate minimizes due to advanced inter-domain flexibility resulting in higher time requirement for preparation of functional complex with the substrate by conformers of the sample which is critical for enzyme catalysis (Kokkinidis et al., 2012). The addition of linkers is a reasonable method for improvement in the level of expression in chimeric proteins. During the cleavage of *in vivo* cleavable linker in chimeric proteins, an improvement was attained in bioactivity, unconventional action, and individual protein domain metabolism (Schmidt, 2013). The bioactivity of chimeric proteins can be improved through increasing the space between fusion proteins by adjusting the length of linkers (Martins et al., 2020).

Strategies for fusion protein design

Interacting recombinant fusion proteins through suitable linkers provide exceptional benefits in correlating domains holding steady complex between unstructured or weak affinity partners. This linker approach does not obstruct the interaction of linked proteins and is beneficial for structural studies. A specific protease cleavage site is usually inserted in the linker region for the efficient removal of a fusion protein (Malik, 2016). Two strategies are used for the rational designing of fusion proteins; one with the help of a suitable linker and another with an *in-silico* approach (Ghavimi et al., 2020). Confirmation of the desired linked protein will be done by characterization as shown in figure 1.

Early research about molecular characteristics of linkers is fueled by computational modeling and crystallography/NMR (Nuclear magnetic resonance) of structured domains, leading to advancement in SAXS (Small-angle X-ray scattering) analyses with data collection (Ruiz et al., 2016a). But these approaches are time-consuming and require expert personnel in comparison with cost-effective and simple computational tools used recently in protein engineering. Several Bioinformatics tools are available for the rational design of chimeric proteins and their linker selection, like Linker DB (Database), LINKER, and SynLinkers (Liu et al., 2015). Sequences of a set of linkers can be automatically generated by LINKER, which is a beneficial tool for fusion protein construction through the generation of linker sequences. Structure Extension With Native-substructure Graphs (SEWING) is a tool that assembles new models from native protein structure based on structural similarity (Ó Conchúir et al., 2015). Prediction of PROtein DOmain boundaries (PPRODO) tool predicts the boundaries of protein domains (Kirillova et al., 2009). Domain linker PRediction using OPTimal features (DROP) differentiate linker and non-linker regions (Ebina et al., 2011). Iterative Threading ASSEMBLY Refinement (I-TASSER) can be used for predicting the 3D structure of a fusion protein (Yang & Zhang, 2015).

Applications of fusion proteins

Recombinant protein construct is widely used to facilitate protein purification and for enhanced expression of soluble proteins in structural biology and biotechnology (Craigo & Feng, 2000). During *in vitro* assays

of proteins, the functional activity can be studied with gene fusion technique (Bushman & Miller, 1997). Bi-functional enzyme engineering can be done with the help of a gene fusion technique.

Fusion genes are employed in cancer diagnostics for the accurate verdict (Maher et al., 2009). Recombinant immunotoxins (RITs) with less immunogenicity are widely used in cancer treatment (Pastan et al., 2006). A toxin fused with an antibody or its fragment can be used to kill a target cancer cell. These toxins can be of bacterial or plant origin and are effectively used in hematological cancers (Hassan et al., 2016). RITs can be used in combination with chemotherapy or for those tumors resistant to chemotherapy like hepatocellular carcinoma (Gao et al., 2015). RITs with an albumin-binding domain (ABD) preferably from streptococcus, targeting mesothelin-expressing cancer cells have been reported to have high anti-tumor activity and a longer half-life in mice models (J. Wei et al., 2018).

Fusion gene can be used to identify the gene regulatory sites in cells, for expression of desired genes, and to manage and measure activity levels of gene regulators. For instance, green fluorescent protein can be observed in tissue or cells through chimeric protein construction of GFP with a gene of interest (Prendergast & Manni, 1978). Fusion protein has a specialized function in structural protein engineering and production (Kim & Pabo, 1997). If gene fusion occurs in the coding sequence, it plays a very important role in the evolution of gene architecture (Eichler, 2001).

In biological traits, variation is created through genetic exchange. Agricultural species can be improved through molecular biology techniques of gene transfer to produce varieties of organisms that show desired characteristics such as a higher proportion of disease-resistant crops (Low et al., 2018). A chimeric protein developed by Monteiro et al demonstrated both antibacterial and antifungal activity. This protein was designed by fusing two select peptides with Blad polypeptide, an active component in a commercially available fungicide; by a flexible linker (Pinheiro et al., 2018).

Industrially useful enzymes can be produced by recombinant technology. Urokinase- plasminogen activator enzyme is produced by a genetically engineered microbe which dissolves blood clot in patients suffering from a heart attack (Adivitiya & Khasa, 2017). Specially developed genetically designed microorganisms are used to clean up the pollutants (Kumar et al., 2013). Gene fusion technique is very helpful in the production of proteins from wastes, enhancement of fermentation processes, and the invention of chemical compounds that have commercial significance (Puetz & Wurm, 2019). A study carried out by Payne, C. M. *et al.* showed that eukaryotic linkers activate glycosylation and enhances binding with cellulose and make them more resistant to proteolysis (Payne et al., 2013).

A small segment of single-stranded DNA that binds with a fluorescent or radioactive marker is used for the diagnosis of some infectious agents like pus-forming *Staphylococcus*, food poisoning *Salmonella*, hepatitis virus, HIV (Human immunodeficiency viruses), etc. Genetic engineering has a significant role in the medical field. Recombinant insulin is produced by genetically engineered bacteria (Baeshen et al., 2014). Antibiotics such as streptomycin and penicillin are produced from genetically prepared *Penicillium sp.* and *Streptomyces sp.* of fungi. Interferon can be produced by recombinant DNA technology ("Interferon Production by Genetic Engineering," 1981).

The selection and production of antibodies can be achieved with this gene fusion technique in biotechnology (R. Arai et al., 2001). Effective linkers can minimize the premature drug release in plasma and advance the selective release in the target cells, thereby enhancing the efficiency of antibody-based drug conjugates (Leung et al., 2020). Fc fusion proteins belong to the recent class of biopharmaceuticals generated by fusing immunoglobulins with biological ligands (Duivelshof et al., 2020).

Conclusion

In this review, we discuss types of linkers that are used for the generation of chimeric protein constructs. Linkers can exert diverse functions and adopt various structures to accomplish the application of fusion proteins. During the last few years, there is an increase in the number of assemblies of chimeric proteins with linkers, which are used for the generation of novel proteins and multifunctional enzymes. The selection of the

linker sequence is very important in fusion protein construction. Challenges in designing of the linker can be met by using appropriate combination and number of amino acids which deliver suitable flexibility with mandatory rigidity such as glycine and serine residues. Designing approaches for suitable linker generation plays an important role to accomplish the desired properties of fusion proteins. The focus should be given to strategies that increase yields and lower costs of target proteins to generate a cost-effective and efficient chimeric protein.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

We thank the Department of Science and Technology for providing INSPIRE Fellowship.

References

- Adivitiya, & Khasa, Y. P. (2017). The evolution of recombinant thrombolytics: Current status and future directions. *Bioengineered* , 8 (4), 331–358. <https://doi.org/10.1080/21655979.2016.1229718>
- Amet, N., Lee, H. F., & Shen, W. C. (2009). Insertion of the designed helical linker led to increased expression of Tf-based fusion proteins. *Pharmaceutical Research* , 26 (3), 523–528. <https://doi.org/10.1007/s11095-008-9767-0>
- Amet, Nurmamet, Wang, W., & Shen, W.-C. (2010). Human growth hormone–transferrin fusion protein for oral delivery in hypophysectomized rats. *Journal of Controlled Release* , 141 (2), 177–182. <https://doi.org/10.1016/j.jconrel.2009.09.007>
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering Design and Selection* , 14 (8), 529–532. <https://doi.org/10.1093/protein/14.8.529>
- Arai, Ryoichi, Wriggers, W., Nishikawa, Y., Nagamune, T., & Fujisawa, T. (2004). *Conformations of Variably Linked Chimeric Proteins Evaluated by Synchrotron X-ray Small-Angle Scattering* . 838 , 829–838. <https://doi.org/10.1002/prot.20244>
- Argos, P. (1990a). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of Molecular Biology* , 211 (4), 943–958. [https://doi.org/10.1016/0022-2836\(90\)90085-Z](https://doi.org/10.1016/0022-2836(90)90085-Z)
- Argos, P. (1990b). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of Molecular Biology* , 211 (4), 943–958. [https://doi.org/10.1016/0022-2836\(90\)90085-Z](https://doi.org/10.1016/0022-2836(90)90085-Z)
- Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., & Redwan, E. M. (2014). Cell factories for insulin production. *Microbial Cell Factories* , 13 (1), 141. <https://doi.org/10.1186/s12934-014-0141-0>
- Bai, Y., & Shen, W. (2006a). *Improving the Oral Efficacy of Recombinant Granulocyte Colony-Stimulating Factor and Transferrin Fusion Protein by Spacer Optimization* . 23 (9), 2116–2121. <https://doi.org/10.1007/s11095-006-9059-5>
- Bai, Y., & Shen, W.-C. (2006b). Improving the oral efficacy of recombinant granulocyte colony-stimulating factor and transferrin fusion protein by spacer optimization. *Pharmaceutical Research* , 23 (9), 2116–2121. <https://doi.org/10.1007/s11095-006-9059-5>
- Bushman, F. D., & Miller, M. D. (1997). *Tethering Human Immunodeficiency Virus Type 1 Preintegration Complexes to Target DNA Promotes Integration at Nearby Sites* . 71 (1), 458–464.

- Chakdar, H., Kumar, M., Pandiyan, K., Singh, A., Nanjappan, K., Kashyap, P. L., & Srivastava, A. K. (2016). Bacterial xylanases: Biology to biotechnology. *3 Biotech* , *6* (2), 1–15. <https://doi.org/10.1007/s13205-016-0457-z>
- Chen, X., Bai, Y., Zaro, J., & Shen, W.-C. (2010a). Design of an in vivo cleavable disulfide linker in recombinant fusion proteins. *BioTechniques* , *49* (1), 513–518. <https://doi.org/10.2144/000113450>
- Chen, X., Bai, Y., Zaro, J., & Shen, W.-C. (2010b). Design of an in vivo cleavable disulfide linker in recombinant fusion proteins. *BioTechniques* , *49* (1), 513–518. <https://doi.org/10.2144/000113450>
- Chen, X., Zaro, J. L., & Shen, W. C. (2013a). Fusion protein linkers: Property, design and functionality. In *Advanced Drug Delivery Reviews* (Vol. 65, Issue 10, pp. 1357–1369). <https://doi.org/10.1016/j.addr.2012.09.039>
- Chen, X., Zaro, J. L., & Shen, W.-C. (2013b). Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* , *65* (10), 1357–1369. <https://doi.org/10.1016/j.addr.2012.09.039>
- Costa, S., Almeida, A., Castro, A., & Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: The novel Fh8 system. *Frontiers in Microbiology* , *5* . <https://doi.org/10.3389/fmicb.2014.00063>
- Crasto, C. J., & Feng, J. (2000). *LINKER : a program to generate linker sequences for fusion proteins* . *13* (5), 309–312.
- Dai, J., Lu, Y., Yu, C., Keller, J. M., Mizokami, A., Zhang, J., & Keller, E. T. (2010). Reversal of chemotherapy-induced leukopenia using GM-CSF promotes bone metastasis that can be blocked with osteoclast inhibitors. *Cancer Research* , *10* (12), 5014–5023. <https://doi.org/10.1111/j.1743-6109.2008.01122.x>. Endothelial
- Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., & Matthews, J. M. (2004). Tandem LIM domains provide synergistic binding in the LMO4:Ldb1 complex. *EMBO Journal* , *23* (18), 3589–3598. <https://doi.org/10.1038/sj.emboj.7600376>
- Dorywalska, M., Dushin, R., Moine, L., Farias, S. E., Zhou, D., Navaratnam, T., Lui, V., Hasa-Moreno, A., Casas, M. G., Tran, T.-T., Delaria, K., Liu, S.-H., Foletti, D., O'Donnell, C. J., Pons, J., Shelton, D. L., Rajpal, A., & Strop, P. (2016). Molecular Basis of Valine-Citrulline-PABC Linker Instability in Site-Specific ADCs and Its Mitigation by Linker Design. *Molecular Cancer Therapeutics* , *15* (5), 958–970. <https://doi.org/10.1158/1535-7163.MCT-15-1004>
- Duivelshof, B. L., Murisier, A., Camperi, J., Fekete, S., Beck, A., Guillaume, D., & D'Atri, V. (2020). Therapeutic Fc-fusion proteins: Current analytical strategies. *Journal of Separation Science* , *jssc.202000765*. <https://doi.org/10.1002/jssc.202000765>
- Ebina, T., Toh, H., & Kuroda, Y. (2011). DROP: An SVM domain linker predictor trained with optimal features selected by random forest. *Bioinformatics* , *27* (4), 487–494. <https://doi.org/10.1093/bioinformatics/btq700>
- Eichler, E. E. (2001). *Recent duplication , domain accretion and the dynamic mutation of the human genome* . *17* (11), 661–669.
- Evers, T. H., Van Dongen, E. M. W. M., Faesen, A. C., Meijer, E. W., & Merckx, M. (2006). Quantitative understanding of the energy transfer between fluorescent proteins connected via flexible peptide linkers. *Biochemistry* , *45* (44), 13183–13192. <https://doi.org/10.1021/bi061288t>
- Gadd, M. S., Bhati, M., Jeffries, C. M., Langley, D. B., Trehwella, J., Guss, J. M., & Matthews, J. M. (2011). Structural basis for partial redundancy in a class of transcription factors, the LIM homeodomain proteins, in neural cell type specification. *Journal of Biological Chemistry* , *286* (50), 42971–42980. <https://doi.org/10.1074/jbc.M111.248559>

- Gao, W., Tang, Z., Zhang, Y.-F., Feng, M., Qian, M., Dimitrov, D. S., & Ho, M. (2015). Immunotoxin targeting glypican-3 regresses liver cancer via dual inhibition of Wnt signalling and protein synthesis. *Nature Communications* , 6 (1), 6536. <https://doi.org/10.1038/ncomms7536>
- George, R A, & Heringa, J. (2003). An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Engineering* , 15 (11), 871–879.
- George, Richard A., & Heringa, J. (2002). An analysis of protein domain linkers: Their classification and role in protein folding. *Protein Engineering* , 15 (11), 871–879. <https://doi.org/10.1093/protein/15.11.871>
- Ghavimi, R., Mohammadi, E., Akbari, V., Shafiee, F., & Jahanian-Najafabadi, A. (2020). In silico design of two novel fusion proteins, p28-IL-24 and p28-M4, targeted to breast cancer cells. *Research in Pharmaceutical Sciences* , 15 (2), 200. <https://doi.org/10.4103/1735-5362.283820>
- Grawe, A., Ranglack, J., Weyrich, A., & Stein, V. (2020). iFLinkC: An iterative functional linker cloning strategy for the combinatorial assembly and recombination of linker peptides with functional domains. *Nucleic Acids Research* , 48 (4), e24–e24. <https://doi.org/10.1093/nar/gkz1210>
- Hassan, R., Alewine, C., & Pastan, I. (2016). New Life for Immunotoxin Cancer Therapy. *Clinical Cancer Research* , 22 (5), 1055–1058. <https://doi.org/10.1158/1078-0432.CCR-15-1623>
- Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., & Crea, R. (1988). Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proceedings of the National Academy of Sciences* , 85 (16), 5879–5883. <https://doi.org/10.1073/pnas.85.16.5879>
- Interferon production by genetic engineering. (1981). *BMJ* ,282 (6265), 674–675. <https://doi.org/10.1136/bmj.282.6265.674-a>
- Kim, J. S., & Pabo, C. O. (1997). Transcriptional repression by zinc finger peptides. Exploring the potential for applications in gene therapy. *J Biol Chem* , 272 (47), 29795–29800. <https://doi.org/#>
- Kingston, R. L., Hamel, D. J., Gay, L. S., Dahlquist, F. W., & Matthews, B. W. (2004). Structural basis for the attachment of a paramyxoviral polymerase to its template. *Proceedings of the National Academy of Sciences* , 101 (22), 8301–8306. <https://doi.org/10.1073/pnas.0402690101>
- Kirillova, S., Kumar, S., & Carugo, O. (2009). Protein Domain Boundary Predictions: A Structural Biology Perspective. *The Open Biochemistry Journal* , 3 , 1–8. <https://doi.org/10.2174/1874091X00903010001>
- Kokkinidis, M., Glykos, N. M., & Fadouloglou, V. E. (2012). Protein flexibility and enzymatic catalysis. In *Advances in Protein Chemistry and Structural Biology* (Vol. 87). <https://doi.org/10.1016/B978-0-12-398312-1.00007-X>
- Kumar, S., Dagar, V. K., Khasa, Y. P., & Kuhad, R. C. (2013). Genetically Modified Microorganisms (GMOs) for Bioremediation. In R. C. Kuhad & A. Singh (Eds.), *Biotechnology for Environmental Management and Resource Recovery* (pp. 191–218). Springer India. https://doi.org/10.1007/978-81-322-0876-1_11
- Leung, D., Wurst, J. M., Liu, T., Martinez, R. M., Datta-Mannan, A., & Feng, Y. (2020). Antibody Conjugates-Recent Advances and Future Innovations. *Antibodies* , 9 (1), 2. <https://doi.org/10.3390/antib9010002>
- Liu, C., Chin, J. X., & Lee, D.-Y. (2015). SynLinker: An integrated system for designing linkers and synthetic fusion proteins: Fig. 1. *Bioinformatics* , 31 (22), 3700–3702. <https://doi.org/10.1093/bioinformatics/btv447>
- Low, L.-Y., Yang, S.-K., Kok, D.-X. A., Ong-Abdullah, J., Tan, N.-P., & Lai, K.-S. (2018). Transgenic Plants: Gene Constructs, Vector and Transformation Method. In O. Celik (Ed.), *New Visions in Plant Science* . InTech. <https://doi.org/10.5772/intechopen.79369>

- Lu, P., & Feng, M. G. (2008a). Bifunctional enhancement of a β -glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Applied Microbiology and Biotechnology* , 79 (4), 579–587. <https://doi.org/10.1007/s00253-008-1468-4>
- Lu, P., & Feng, M. G. (2008b). Bifunctional enhancement of a β -glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Applied Microbiology and Biotechnology* , 79 (4), 579–587. <https://doi.org/10.1007/s00253-008-1468-4>
- Maher, C. A., Palanisamy, N., Brenner, J. C., Cao, X., Kalyana-Sundaram, S., Luo, S., Khrebtukova, I., Barrette, T. R., Grasso, C., Yu, J., Lonigro, R. J., Schroth, G., Kumar-Sinha, C., & Chinnaiyan, A. M. (2009). Chimeric transcript discovery by paired-end transcriptome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* , 106 (30), 12353–12358. <https://doi.org/10.1073/pnas.0904720106>
- Malik, A. (2016). Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. *3 Biotech* , 6 (1), 44. <https://doi.org/10.1007/s13205-016-0397-7>
- Martins, M., Dinamarco, T. M., & Goldbeck, R. (2020). Recombinant chimeric enzymes for lignocellulosic biomass hydrolysis. *Enzyme and Microbial Technology* , 140 , 109647. <https://doi.org/10.1016/j.enzmictec.2020.109647>
- McCormick, A. L., Thomas, M. S., & Heath, A. W. (2001). Immunization with an interferon-gamma-gp120 fusion protein induces enhanced immune responses to human immunodeficiency virus gp120. *J Infect Dis* , 184 (11), 1423–1430. <https://doi.org/10.1086/324371>
- Nagamune, T. (2017). Biomolecular engineering for nanobio/bionanotechnology. *Nano Convergence* , 4 (1), 9. <https://doi.org/10.1186/s40580-017-0103-4>
- Nick Pace, C., & Martin Scholtz, J. (1998). A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophysical Journal* , 75 (1), 422–427. [https://doi.org/10.1016/S0006-3495\(98\)77529-0](https://doi.org/10.1016/S0006-3495(98)77529-0)
- Ó Conchúir, S., Barlow, K. A., Pache, R. A., Ollikainen, N., Kundert, K., O’Meara, M. J., Smith, C. A., & Kortemme, T. (2015). A Web Resource for Standardized Benchmark Datasets, Metrics, and Rosetta Protocols for Macromolecular Modeling and Design. *PLOS ONE* , 10 (9), e0130433. <https://doi.org/10.1371/journal.pone.0130433>
- Pace, C. N., & Scholtz, J. M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* , 75 (1), 422–427. [https://doi.org/10.1016/S0006-3495\(98\)77529-0](https://doi.org/10.1016/S0006-3495(98)77529-0)
- Pastan, I., Hassan, R., FitzGerald, D. J., & Kreitman, R. J. (2006). Immunotoxin therapy of cancer. *Nature Reviews Cancer* , 6 (7), 559–565. <https://doi.org/10.1038/nrc1891>
- Payne, C. M., Resch, M. G., Chen, L., Crowley, M. F., Himmel, M. E., Taylor, L. E., Sandgren, M., Ståhlberg, J., Stals, I., Tan, Z., & Beckham, G. T. (2013). Glycosylated linkers in multimodular lignocellulose-degrading enzymes dynamically bind to cellulose. *Proceedings of the National Academy of Sciences* , 110 (36), 14646–14651. <https://doi.org/10.1073/pnas.1309106110>
- Pinheiro, A. M., Carreira, A., Ferreira, R. B., & Monteiro, S. (2018). Fusion proteins towards fungi and bacteria in plant protection. *Microbiology* , 164 (1), 11–19. <https://doi.org/10.1099/mic.0.000592>
- Poreba, M. (2020). Protease-activated prodrugs: Strategies, challenges, and future directions. *The FEBS Journal* , 287 (10), 1936–1969. <https://doi.org/10.1111/febs.15227>
- Prendergast, F. G., & Manni, K. G. (1978). *Chemical and Physical Properties of Aequorin and the Green Fluorescent Protein Isolated from Aequorea forskilea* + . 3448–3453.
- Priyanka, V., Chichili, R., Kumar, V., & Sivaraman, J. (2013a). *Linkers in the structural biology of protein – protein interactions* (Vol. 22, pp. 153–167). <https://doi.org/10.1002/pro.2206>

- Priyanka, V., Chichili, R., Kumar, V., & Sivaraman, J. (2013b). *Linkers in the structural biology of protein – protein interactions* . 22 , 153–167. <https://doi.org/10.1002/pro.2206>
- Puetz, J., & Wurm, F. M. (2019). Recombinant Proteins for Industrial versus Pharmaceutical Purposes: A Review of Process and Pricing. *Processes* , 7 (8), 476. <https://doi.org/10.3390/pr7080476>
- Receveur, V., Czjzek, M., Schulein, M., Panine, P., & Henrissat, B. (2002a). Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle X-ray scattering. *Journal of Biological Chemistry* , 277 (43), 40887–40892. <https://doi.org/10.1074/jbc.M205404200>
- Receveur, V., Czjzek, M., Schulein, M., Panine, P., & Henrissat, B. (2002b). Dimension, shape, and conformational flexibility of a two domain fungal cellulase in solution probed by small angle X-ray scattering. *Journal of Biological Chemistry* , 277 (43), 40887–40892. <https://doi.org/10.1074/jbc.M205404200>
- Reddy Chichili, V. P., Kumar, V., & Sivaraman, J. (2013). Linkers in the structural biology of protein-protein interactions: Linkers in Structural Biology. *Protein Science* , 22 (2), 153–167. <https://doi.org/10.1002/pro.2206>
- Robinson, C. R., & Sauer, R. T. (1998a). Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Computational Biomolecular Science* , 95 (May), 5929–5934. <https://doi.org/10.1073/pnas.95.11.5929>
- Robinson, C. R., & Sauer, R. T. (1998b). Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Computational Biomolecular Science* , 95 (May), 5929–5934. <https://doi.org/10.1073/pnas.95.11.5929>
- Rożycki, B., Cazade, P.-A., O’Mahony, S., Thompson, D., & Cieplak, M. (2017). <https://doi.org/10.1039/R5CP004114D>. *Physical Chemistry Chemical Physics* , 19 (32), 21414–21425. <https://doi.org/10.1039/R5CP004114D>
- Ruiz, D. M., Turowski, V. R., & Murakami, M. T. (2016a). Effects of the linker region on the structure and function of modular GH5 cellulases. *Scientific Reports* , 6 (March), 1–13. <https://doi.org/10.1038/srep28504>
- Ruiz, D. M., Turowski, V. R., & Murakami, M. T. (2016b). Effects of the linker region on the structure and function of modular GH5 cellulases. *Scientific Reports* , 6 , 28504. <https://doi.org/10.1038/srep28504>
- Saadat, F. (2017). A review on chimeric xylanases: Methods and conditions. *3 Biotech* , 7 (1), 67. <https://doi.org/10.1007/s13205-017-0660-6>
- Sammond, D. W., Payne, C. M., Brunecky, R., Himmel, M. E., Crowley, M. F., & Beckham, G. T. (2012). Cellulase linkers are optimized based on domain type and function: Insights from sequence analysis, biophysical measurements, and molecular simulation. *PloS One* , 7 (11). <https://doi.org/10.1371/journal.pone.0048615>
- Schmidt, S. R. (2013). *Fusion Protein Technologies for Biopharmaceuticals: Applications and Challenges PART I INTRODUCTION* . <https://doi.org/10.13140/2.1.4445.5366>
- Tollefsen, S., Hotta, K., Chen, X., Simonsen, B., Swaminathan, K., Mathews, I. I., Sollid, L. M., & Kim, C. Y. (2012). Structural and functional studies of trans-encoded HLA-DQ2.3 (DQA1*03:01/DQB1*02:01) protein molecule. *Journal of Biological Chemistry* , 287 (17), 13611–13619. <https://doi.org/10.1074/jbc.M111.320374>
- Tsuchikama, K., & An, Z. (2018). Antibody-drug conjugates: Recent advances in conjugation and linker chemistries. *Protein & Cell* , 9 (1), 33–46. <https://doi.org/10.1007/s13238-016-0323-0>
- Tuna, M., Amos, C. I., & Mills, G. B. (2019). Molecular mechanisms and pathobiology of oncogenic fusion transcripts in epithelial tumors. *Oncotarget* , 10 (21), 2095–2111. <https://doi.org/10.18632/oncotarget.26777>

van Rosmalen, M., Krom, M., & Merckx, M. (2017). Tuning the Flexibility of Glycine-Serine Linkers To Allow Rational Design of Multidomain Proteins. *Biochemistry*, *56* (50), 6565–6574. <https://doi.org/10.1021/acs.biochem.7b00902>

Wei, B., Gunzner-Toste, J., Yao, H., Wang, T., Wang, J., Xu, Z., Chen, J., Wai, J., Nonomiya, J., Tsai, S. P., Chuh, J., Kozak, K. R., Liu, Y., Yu, S.-F., Lau, J., Li, G., Phillips, G. D., Leipold, D., Kamath, A., ... Safina, B. (2018). Discovery of Peptidomimetic Antibody–Drug Conjugate Linkers with Enhanced Protease Specificity. *Journal of Medicinal Chemistry*, *61* (3), 989–1000. <https://doi.org/10.1021/acs.jmedchem.7b01430>

Wei, J., Bera, T. K., Liu, X. F., Zhou, Q., Onda, M., Ho, M., Tai, C.-H., & Pastan, I. (2018). Recombinant immunotoxins with albumin-binding domains have long half-lives and high antitumor activity. *Proceedings of the National Academy of Sciences*, *115* (15), E3501–E3508. <https://doi.org/10.1073/pnas.1721780115>

Werner, J. H., Joggerst, R., Dyer, R. B., & Goodwin, P. M. (2006). A two-dimensional view of the folding energy landscape of cytochrome c. *Proceedings of the National Academy of Sciences of the United States of America*, *103* (30), 11130–11135. <https://doi.org/10.1073/pnas.0604712103>

Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, *43* (W1), W174–W181. <https://doi.org/10.1093/nar/gkv342>

Zhao, H. L., Yao, X. Q., Xue, C., Wang, Y., Xiong, X. H., & Liu, Z. M. (2008a). Increasing the homogeneity, stability and activity of human serum albumin and interferon- α 2b fusion protein by linker engineering. *Protein Expression and Purification*, *61* (1), 73–77. <https://doi.org/10.1016/j.pep.2008.04.013>

Zhao, H. L., Yao, X. Q., Xue, C., Wang, Y., Xiong, X. H., & Liu, Z. M. (2008b). Increasing the homogeneity, stability and activity of human serum albumin and interferon- α 2b fusion protein by linker engineering. *Protein Expression and Purification*, *61* (1), 73–77. <https://doi.org/10.1016/j.pep.2008.04.013>

Table 1: Summary of linker types

Types of Linker	Features	Examples	Advantages	Applications
Flexible	Preservation of cooperative inter-domain interactions. Delivers flexibility and mobility to linking functional domains	HSA-IFN- α 2b(Zhao et al., 2008b)	Maintain space among functional domains. Interaction or movement with a certain degree for joined domains	Improvement in activity and stability of human serum albumin and interferon-alpha2b chimera.
Rigid	Proline residues with α - helical structures. Generate relatively stable structures	hGH-Tf(Nurmamet Amet et al., 2010)	Powerfully separate functional domains. Preserve bioactivity of chimeric proteins	In oral delivery of human growth hormone
<i>In vivo</i> Cleavable	Perform as single-molecule and cleave functional domains during the <i>in-vivo</i> process.	G-CSF-S-S-Tf(Chen et al., 2010b)	Enable targeting. Achieve optimal biological activity.	Enhances the therapeutic effect.

List of Figures

Figure 1. Linker strategy for chimeric protein construction

Figure 2. A diagrammatic representation of different types of linkers

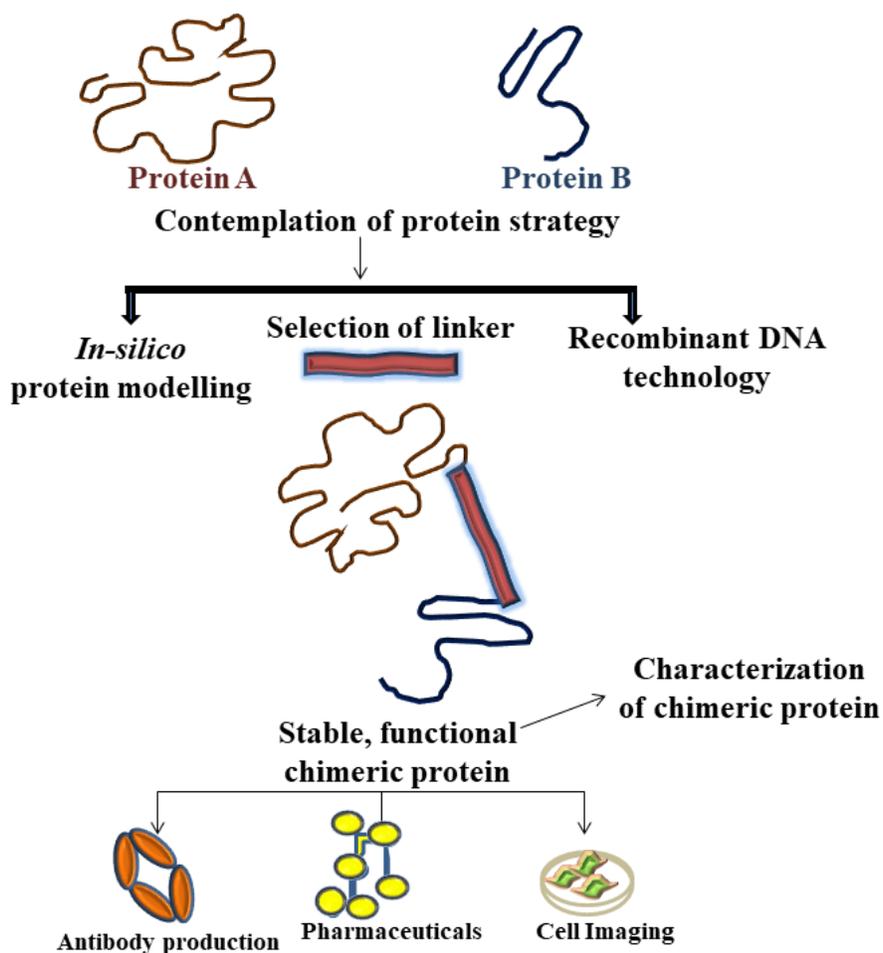
Figure 3(a). Flexible linker showing a loop-like structure. (Gly-Gly-Gly-Gly-ser)_n present between V_H and V_L chain of the antibody

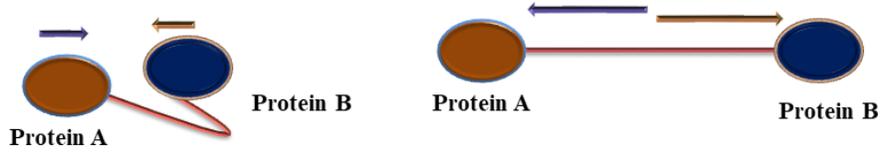
Figure 3(b). Rigid linker showing an α -helical structure

Figure 3(c). *In-vivo* cleavable linker

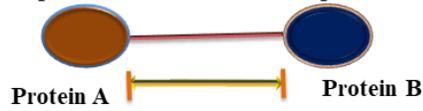
Hosted file

table.docx available at <https://authorea.com/users/364767/articles/485067-linkers-a-synergistic-way-for-chimeric-proteins>





Flexible linker: Delivers flexibility and mobility to linking functional domains which permits interaction and separation between domains



Rigid linker: Brings stability and maintain distance between functional domains for stable structure generation



***In-vivo* cleavable linker:** Cleaves functional domains during *in-vivo* process

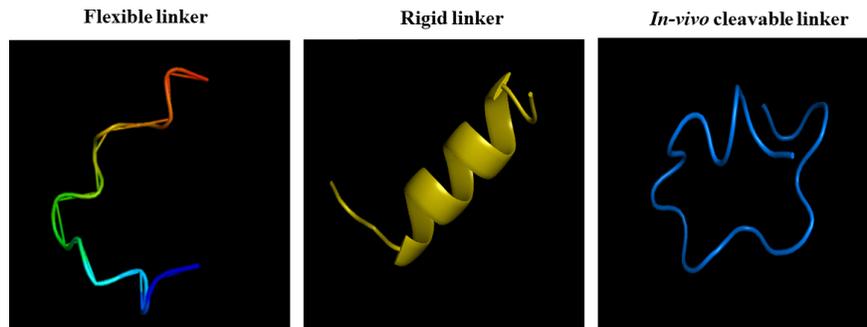


Figure 3(a)

Figure 3(b)

Figure 3(c)